

**Universität
Rostock**



Traditio et Innovatio

Statistische Analyse von Aktinfilamenten und Anpassungstests mit geschätzten Parametern

Dissertation zur Erlangung des akademischen Grades

doctor rerum naturalium (Dr. rer. nat.)

der Mathematisch-Naturwissenschaftlichen Fakultät

der Universität Rostock

vorgelegt von

Bing Liu, geb. am 07.03.1981 in LiaoNing, China

aus Hamburg

Betreuer: Prof. Dr. Friedrich Liese

Rostock, den 30. Mai 2012

urn:nbn:de:gbv:28-diss2013-0007-0

Gutachter:

Prof. Dr. rer. nat. habil. Friedrich Liese

Mathematische Statistik mit Schwerpunkt stochastische Prozesse

Universität Rostock

Prof. Dr. rer. nat. habil. Silvelyn Zwanzig

Mathematische Statistik

Universität Uppsala, Schweden

Verteidigungsdatum: 18. Oktober 2012

Inhaltsverzeichnis

1	Einleitung	6
1.1	Statistische Analyse von Aktinfilamenten	6
1.2	Anpassungstests mit geschätzten Parametern	9
2	Modellierung von Aktinfilamenten	12
2.1	Der Streckenprozess der Aktinfilamente	12
2.2	Existenz von Hauptrichtungen	14
2.3	Bivariate homogene Verteilungen	18
2.4	Bivariate elliptisch konturierte Verteilungen	28
3	Testen auf Hauptrichtungen des Wachstums	35
3.1	Testen in normalverteilten Grundgesamtheiten	35
3.1.1	Verteilung der empirischen Kovarianzmatrix	35
3.1.2	Invariante Teststatistiken	38
3.1.3	Likelihoodquotientest in normalverteilten Grundgesamtheiten	39
3.2	Asymptotische Tests auf Vorliegen einer Hauptrichtung im Einclustermodell . .	47
3.3	Asymptotische Tests auf Vorliegen einer Hauptrichtung im Mehrclustermodell .	54
3.4	Bedingte Tests auf Vorliegen einer Hauptrichtung	59
3.5	Computersimulationen	62
3.5.1	Niveau der Tests	62
3.5.2	Die asymptotische Güte unter speziellen Alternativen	64
3.6	Test von Projektdaten	65
4	Anpassungstests	69
4.1	Grenzwertsätze für die empirische Verteilungsfunktion	69
4.1.1	Konvergenz unter einfacher Nullhypothese	69
4.1.2	Konvergenz unter lokalen Alternativen	70
4.1.3	Konvergenz des empirischen Prozesses mit geschätzten Parametern . .	75
4.2	Der randomisierte empirische Prozesses	78
4.3	Der Cramer-von-Mises-Test für spezielle Modelle und spezielle Alternativen . .	82
4.3.1	Normalverteilung mit unbekanntem μ und bekanntem σ^2	82
4.3.2	Normalverteilung mit unbekannten μ und σ^2	87
4.3.3	Exponentialverteilung mit unbekanntem λ	89
4.4	Computersimulationen	90
4.4.1	Niveau der Anpassungstests	90
4.4.2	Die asymptotische Güte unter speziellen Alternativen	91
5	Anhang	94
5.1	Anhang I: Die Lemmata von LeCam	94
5.2	Anhang II: Schwache Konvergenz im Skorochod-Raum	97
5.3	Anhang III: Brownsche Brücke und Kac-Siebert-Zerlegung	98
5.4	Anhang IV: Spezielle Grenzwertsätze	100
5.5	Anhang V: Quadratische Formen von normalverteilten Zufallsvariablen	101
6	Literaturverzeichnis	103

Inhaltsverzeichnis

7	Zusammenfassung	106
8	Selbständigkeitserklärung	107

1 Einleitung

Die vorliegende Dissertation entstand im Rahmen des Graduiertenkollegs Welisa. Wie bei allen Problemen der angewandten Statistik ist die Wahl eines geeigneten Modells von entscheidender Bedeutung. Deshalb wurde bereits in einem frühen Stadium der Dissertation dieses Problem als eigenständiges mathematisches Problem behandelt. Die wesentlichen Ergebnisse dieser Untersuchung sind im Kapitel 4 dieser Arbeit dargestellt und in Liese und Liu (2011) veröffentlicht. Die mit dem Graduiertenkolleg unmittelbar zusammenhängenden Probleme betreffen die mathematischen Modelle und Analysen für Aktinfilamente. Die Ergebnisse sind in den Kapiteln 2 und 3 zusammengefasst.

1.1 Statistische Analyse von Aktinfilamenten

Aktin ist ein Strukturprotein und Bestandteil des Zytoskeletts. In der Zelle bildet Aktin dynamische Filamente, die sogenannten Aktinfilamente. Diese Filamente dienen als Bestandteil des Zytoskeletts der Stabilisierung der äußeren Zellform¹. Ausgangspunkt für die statistische Modellierung von Aktinfilamenten sind Bilder von Aktinfilamenten, die mit Hilfe der Gratzmeyer-Methode als Methode zur Bildauswertung in der Dissertation von Herrn Birkholz² gewonnen wurden. Diese Bilder zeigen, dass sich in Abhängigkeit von der Oberflächenstruktur auf der die Zellen kultiviert wurden, Vorzugsrichtungen herausbilden, entlang derer die Filamente besonders lang sind. Der statistische Nachweis einer solchen Hauptrichtung benötigt ein statistisches Modell innerhalb dessen dann entsprechende Tests konstruiert werden können. Ein erster Modellansatz basiert auf der Tatsache, dass Punkte aus einem Regressionsmodell sich um die Regressionsgerade gruppieren. Verbindet man Paare von Punkten, so erhält man zufällige Strecken, die zueinander mit einer gewissen Parallelität verlaufen. Simulationen für Punkte aus einem Regressionsmodell zeigen, dass die Parallelität groß ist und damit eine gewisse Vorzugsrichtung gesichert ist, wenn sich die beiden Eigenwerte der Kovarianzmatrix stark unterscheiden. Damit läuft der Nachweis einer Hauptrichtung für das Wachstum von Filamenten auf einen Test für die Gleichheit von Eigenwerten der Kovarianzmatrix hinaus. Diese Abhängigkeit der Existenz einer Hauptrichtung von den Eigenwerten wird auch klar, wenn man den Argumenten der Hauptkomponentenanalyse folgt. Fragt man danach, ob es eine Richtung in der Ebene gibt, für die die Komponente des Filaments größer als auf der dazu senkrecht stehenden Richtung ist, so führt dies wieder auf den Nachweis, dass sich die beiden Eigenwerte der Kovarianzmatrix unterscheiden. Solche Tests beantworten aber noch nicht die Frage, wie die Länge mit dem Winkel der Filamente zusammenhängt. Hierzu sind geeignete zweidimensionale Verteilungen notwendig. Liegt eine zweidimensionale Normalverteilung vor, so kann man durch Übergang zu Polarkoordinaten leicht die bedingte Verteilung der Länge bei gegebenem Winkel und den zugehörigen bedingten Erwartungswert formelmäßig berechnen. Es zeigt sich nun, dass dieser bedingte Erwartungswert der Länge genau dann vom Winkel unabhängig ist, wenn beide Eigenwerte übereinstimmen. Darüber hinaus gilt dies genau dann, wenn der Winkel und die Länge voneinander unabhängig sind. Um eine große Klasse von zweidimensionalen Verteilungen für die Modellierung von Aktinfilamenten bereitzustellen, wurde das von Hoffmann-Jørgensen (1994) eingeführte Konzept der homogenen Dichten aufgegriffen. Diese Verteilungsklasse enthält die Normalverteilung, die elliptisch konturierten

¹Aktin: Wikipedia: <http://de.wikipedia.org/wiki/Aktin>

²Dissertation: Mathematical Methods for the Quantification of Actin-Filaments in Microscopic Images, 2012

1 Einleitung

Verteilungen und die $I_{n,p}$ -symmetrischen Verteilungen als Spezialfälle. Einerseits ist diese Verteilungsklasse sehr flexibel. Andererseits ist sie besonders gut geeignet, die Zusammenhänge zwischen der Länge und dem Winkel eines zweidimensionalen Vektors in einfacher Weise zu beschreiben. Insbesondere liegt wieder eine Unabhängigkeit des Winkels von der Länge vor, wenn man zu einer modifizierten Länge übergeht. Diese stochastische Unabhängigkeit führt zu einer stochastischen Darstellung eines zufälligen Vektors mit homogener Dichte. In diese Darstellung gehen der Winkel und die modifizierte Länge ein, die durch einen Punkt auf der Konstanzlinie der Dichte bestimmt ist. Im Falle einer Normalverteilung sind diese Konstanzlinien Ellipsen, die bei unkorrelierten Komponenten des zufälligen Vektors in Kreise übergehen. Gegenstand des zweiten Kapitel dieser Dissertation ist die Bereitstellung der wahrscheinlichkeitstheoretischen Zusammenhänge in der Klasse der homogenen Dichten. Dies betrifft Aussagen über die Verteilungen der Längen, der Winkel, der beiden bedingten Dichten und bedingten Erwartungswerte. Einige dieser Resultate sind bereits in Hoffmann-Jørgensen (1994) zu finden, werden hier aber nochmals direkt mit Hilfe von Polarkoordinaten und ohne Verwendung der in Hoffmann-Jørgensen (1994) verwendeten Hilfsmittel der geometrischen Maßtheorie abgeleitet. Im Kapitel 3 werden statistische Tests konstruiert, mit deren Hilfe nachgewiesen werden soll, ob es Vorzugsrichtungen für die Länge der Filamente gibt. Im Kapitel 2 wird herausgearbeitet, dass diese Frage darauf hinausläuft, ob die Eigenwerte der Kovarianzmatrix gleich sind. Zu dieser klassischen Fragestellung gibt es für normalverteilte Grundgesamtheiten beginnend mit Mauchly (1940a) und (1940b) zahlreiche Arbeiten, die auch den Fall einer allgemeinen Dimension erfassen. Folgt man dem klassischen Zugang in Anderson (1984) bei Beschränkung auf drehungsinvariante Statistiken, so reduzieren sich die möglichen Teststatistiken auf Funktionen der Eigenwerte $L_{n,1}, \dots, L_{n,d}$ der empirischen Kovarianzmatrix, die aus der Beobachtung von n unabhängigen d -dimensionalen Vektoren berechnet werden. Unter den verschiedenen Teststatistiken zum Testen der Gleichheit der Eigenwerte $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_d$, die auf Funktionen der $L_{n,1}, \dots, L_{n,d}$ beruhen, ergibt sich in normalverteilten Grundgesamtheiten die Teststatistik

$$L_n = [L_{n,1} \cdot \dots \cdot L_{n,d}]^{1/d} \left[\frac{1}{d} \sum_{i=1}^d L_{n,i} \right]^{-1}$$

als Likelihoodquotientenstatistik. Die Verteilung dieser Teststatistik unter der Nullhypothese $\lambda_1 = \lambda_2 = \dots = \lambda_d$ ist ein klassisches Resultat der multivariaten Statistik. Für die Herleitung der Verteilung von L_n unter der Alternative benötigt man die Verteilung des Vektors $(L_{n,1}, \dots, L_{n,d})$ unter der Alternative. Dann tritt in der Dichteformel ein Integral über die orthogonale Gruppe auf. Dieses Integral lässt sich nicht geschlossen berechnen, siehe Muirhead (1982). Da wir nur den Fall der Dimension $d = 2$ betrachten, können wir diese Schwierigkeit umgehen, indem wir auf Girshick (1941) für eine direkte Herleitung der Verteilung von L unter der Alternativen zurückgreifen. Der in Girshick (1941) gegebene Beweis ist an einigen Stellen stark verkürzt und enthält Lücken. Deshalb wird im Kapitel 3 ein ausführlicher Beweis gegeben, wobei wir aber wesentliche Teile des Girshickschen Beweises verwenden. Mit Hilfe der Verteilung von L_n unter der Alternative lässt sich die Gütefunktion des Mauchlytests berechnen. Liegt keine Normalverteilung vor, so ist man auf eine asymptotische Approximation der Verteilung der Teststatistik angewiesen. Das asymptotische Verhalten der Folge $T_n = n(1 - L_n^2)$ wurde für beliebige Dimension bereits in Waternaux (1977) für den Fall elliptisch konturierter Verteilungen untersucht. Dort wird auch die asymptotische Verteilung unter lokalen Alternativen abgeleitet. Für den allgemeinen Fall, wo nur die Endlichkeit gewisser Momente gefordert wird, wurden Grenzwertsätzen unter der Nullhypothese in Gupta und Xu

1 Einleitung

(2006) und in der dort zitierten Literatur abgeleitet. Als Grenzverteilungen treten Verteilungen von Linearkombinationen χ^2 -verteilter Zufallsvariablen auf. Dieser Effekt des Abweichens der Grenzverteilung von der χ^2 -Verteilung wird in der Literatur über multivariate Statistik als Nichtrobustheit bezeichnet. Wir verallgemeinern für die Dimension zwei die Ergebnisse von Waternaux (1977) und Gupta und Xu (2006). Weiterhin modifizieren wir die Teststatistik $T_n = n(1 - L_n^2)$ so, dass die neue Teststatistik Q_n in dem oben geschilderten Sinne robust ist. Die zunächst für ein Dreiecksschema bewiesenen Resultate werden im zweiten Abschnitt von Kapitel 3 angewendet, um festzustellen, ob sich durch Versuchsänderungen neue Haupttrichtungen herausgeprägt haben. Wir nehmen an, dass die beobachteten Daten die Verteilung $\mathcal{L}(A(X, Y)^T)$ haben und wollen überprüfen, ob A weder aus einer reinen Streckung noch aus einer Drehung besteht. Das ist gleichbedeutend damit dass AA^T gleiche Eigenwerte besitzt. Es sei Q_A die zu

$$f_A(x, y) = (\det(A))^{-1} f(A^{-1}(x, y)^T)$$

gehörige Verteilung, wobei f den Erwartungswert Null und eine Diagonalmatrix als Kovarianzmatrix besitzt. Das führt auf das folgende Testproblem

$$\text{Nullhypothese: } H_0 : \Sigma = AA^T = \sigma^2 I$$

$$\text{Alternativhypothese: } H_A : \Sigma = AA^T \neq \sigma^2 I.$$

Die Güte der betrachteten asymptotischen Tests wird unter zwei speziellen Alternativen untersucht. Weil die Teststatistik Q_n noch eine aus gemischten vierten Momenten bestehende Matrix Γ enthält, muss diese durch einen nichtparametrischen Schätzer $\hat{\Gamma}_n$ mitgeschätzt werden. Weil Γ multiplikativ in Q_n eingeht, verändert sich dadurch die Asymptotik der neuen Teststatistik \hat{Q}_n gegenüber Q_n nicht. Mit Hilfe der Teststatistiken T_n und \hat{Q}_n werden asymptotische α -Tests konstruiert, deren asymptotische Güte unter den lokalen Alternativen $H_{A,loc}$ berechnet wird. In diese Gütefunktionen geht die Matrix Γ ein, wodurch der Einfluss von schweren Flanken der Verteilungen auf die Güte erfasst wird.

Im dritten Abschnitt von Kapitel 3 wird ein asymptotischer Test konstruiert, wenn die beobachteten zufälligen Vektoren $(X_1, Y_1), \dots, (X_n, Y_n)$ nicht notwendig identisch normalverteilt sind, aber den Erwartungswert Null haben und Kovarianzmatrizen besitzen, deren Eigenwerte λ_1, λ_2 identisch sind. Genauer möge es gewisse orthogonalen Matrizen O_j derart geben, dass die Vektoren $(U_j, V_j)^T = O_j^T (X_j, Y_j)^T$ unabhängig sind und die gleiche Normalverteilung besitzen. Dies entspricht einer Situation, wenn die Filamente aus verschiedenen Clustern stammen, die durch Translationen und Drehungen ineinander überführt werden können. Geht man jetzt zu den Längenquadraten

$$R_j^2 = U_j^2 + V_j^2 = X_j^2 + Y_j^2$$

über, dann liegen unabhängige und identisch verteilte Beobachtungen vor. Die Verteilung von R_j^2 hängt von λ_1, λ_2 ab. Durch Vertauschen von λ_1 und λ_2 ergibt sich die gleiche Verteilung. Damit sind die Parameter nicht identifizierbar. Schränkt man den Parameterbereich durch die Forderung $\lambda_1 \geq \lambda_2$ ein, so liegt die Nullhypothese am Rand des Parameterbereichs, was in den klassischen Sätzen über den Likelihoodquotiententest ausgeschlossen ist. Deshalb wird die auf Momenten basierte Teststatistik

$$T_n = \frac{(\frac{1}{n} \sum_{i=1}^n R_i^2)^2}{\frac{1}{n} \sum_{i=1}^n R_i^4}.$$

verwendet und deren asymptotisches Verhalten untersucht. Es wird gezeigt, dass unter der Normalverteilungsannahme die auf den T_n basierende Folge von α -Tests global konsistent ist.

1 Einleitung

Die Schwierigkeit der asymptotischen Tests für die Überprüfung der Gleichheit der Eigenwerte besteht darin, dass die Fehlerwahrscheinlichkeit erster Art nicht exakt eingehalten wird. Bestehen unter der Nullhypothese die Daten aus zwei unabhängigen Bestandteilen, so lassen sich bedingte Tests konstruieren, die dann exakte Tests in dem Sinne sind, dass sie das Niveau exakt einhalten. Sind die Daten eindimensional, so sind solche unabhängigen Bestandteile die Rangstatistik und die Ordnungsstatistik. Für homogene Verteilungen sind die modifizierte Länge und der Winkel solche unabhängigen Bestandteile, mit deren Hilfe bedingte Tests konstruiert werden. Die Bestimmung der kritischen Werte für die bedingten Verteilungen kann durch Simulationen erfolgen, weil unter der Nullhypothese die bedingten Verteilungen die nichtparametrische Komponente der unbekannten Längenverteilung nicht mehr enthalten.

1.2 Anpassungstests mit geschätzten Parametern

Die Wahl eines geeigneten parametrischen statistischen Modells gehört zu den grundlegenden Aufgaben der mathematischen Statistik. Dieses Problem wird im Kapitel 4 der vorliegenden Dissertation untersucht. Die klassische Fragestellung im Rahmen der Anpassungstests betrifft folgendes Testproblem. Es seien X_1, \dots, X_n u.i.v. Zufallsvariablen mit reellen Werten und der Verteilung $\mathcal{L}(X_1)$. Weiterhin sei P eine vorgegebene Verteilung auf der σ -Algebra der Borelmengen \mathfrak{B} der reellen Achse $\mathbb{R} = (-\infty, \infty)$. Getestet werden soll die Nullhypothese $H_0 : \mathcal{L}(X_1) = P$ gegen die Alternative $H_A : \mathcal{L}(X_1) \neq P$. Zu diesem Zweck wird zunächst die empirische Verteilungsfunktion

$$\hat{F}_n(t) = \frac{1}{n} \sum_{i=1}^n I_{(-\infty, t]}(X_i)$$

gebildet. Sei $F(t) = P((-\infty, t])$ die Verteilungsfunktion von P . Nach dem Satz von Glivenko-Cantelli gilt fast sicher

$$\lim_{n \rightarrow \infty} \sup_{-\infty < t < \infty} |\hat{F}_n(t) - F(t)| = 0.$$

Deshalb bauen viele Tests auf einen Vergleich von $\hat{F}_n(t)$ mit $F(t)$ mit Hilfe geeigneter Funktionale auf. Beispielhaft sei hier der Kolmogoroff-Smirnov-Test genannt, der auf der Tatsache beruht, dass unter der Nullhypothese $\mathcal{L}(X_1) = P$ die Verteilungen von

$$K_n = \sup_{-\infty < t < \infty} \sqrt{n} |\hat{F}_n(t) - F(t)|$$

gegen eine Grenzverteilung streben, die von F unabhängig ist, falls F stetig ist. Das Testen einer einfachen Nullhypothese betrifft nicht die Mehrzahl der in der Praxis auftretenden Anpassungsprobleme. Sei $(P_\theta)_{\theta \in \Theta}$, $\Theta \subseteq \mathbb{R}^d$, eine parametrisierte Familie von Verteilungen. Es soll festgestellt werden, ob die Verteilung $\mathcal{L}(X_1)$ zu $(P_\theta)_{\theta \in \Theta}$ gehört oder nicht. Wir betrachten also das Testproblem

$$H_0 : \mathcal{L}(X_1) \in (P_\theta)_{\theta \in \Theta} \quad \text{gegen} \quad H_A : \mathcal{L}(X_1) \notin (P_\theta)_{\theta \in \Theta}.$$

Sei $F_\theta(t) = P_\theta((-\infty, t])$. Eine nahe liegende Idee besteht darin, den unbekannten Parameter θ durch eine Schätzung $\hat{\theta}_n$ zu ersetzen und damit die Teststatistik

$$\tilde{K}_n = \sup_{-\infty < t < \infty} \sqrt{n} |\hat{F}_n(t) - F_{\hat{\theta}_n}(t)|$$

1 Einleitung

zu erhalten. Es zeigt sich, dass bei schwachen Regularitätsvoraussetzungen unter der Nullhypothese die Verteilungen dieser Teststatistiken gegen eine Grenzverteilung streben, die aber jetzt von der Verteilungsfamilie $(P_\theta)_{\theta \in \Theta}$ und dem unbekannten Parameter θ abhängt. Damit sind die betrachteten Tests nicht mehr asymptotisch verteilungsfrei. In einem wichtigen Spezialfall kann man diese Schwierigkeiten teilweise umgehen. Für die Skalen-Lokationsfamilie

$$F_\theta(t) = F\left(\frac{t - \mu}{\sigma}\right), \theta = (\mu, \sigma), \sigma > 0, \mu \in \mathbb{R}$$

hängt die Grenzverteilung von \tilde{K}_n nur noch von F aber nicht mehr von θ ab. Eine weitere Methode zur Umgehung der oben beschriebenen Schwierigkeiten bei Anpassungstests mit geschätzten Parametern besteht darin, spezielle Funktionale von Verteilungen zu verwenden, die innerhalb einer gegebenen Klasse von Verteilungen einen bestimmten Wert, etwa Null haben. Für die Klasse der Normalverteilungen sind hier die auf der Schiefe und dem Exzess beruhenden Tests zu nennen, zu denen insbesondere der Bowman-Shenton-Test (auch Jarque-Bera-Test genannt) gehört. Ein Nachteil dieser Tests besteht darin, dass sie nicht notwendig global konsistent sind. Es gab bis Ende der 70'er Jahre keine systematische Theorie der Eliminierung des störenden Parameters θ , um die Anpassungstests asymptotisch verteilungsfrei zu machen. Einen Durchbruch erzielte Khmaladze (1981), der einen Kern K_θ so konstruierte, dass

$$\hat{B}_n(t) - \sqrt{n} \int \left(\int K_\theta(s, t, \tau) P_\theta(ds) \right) d(\hat{F}_n(\tau) - F_{\hat{\theta}_n}(\tau))$$

gegen einen Prozeß strebt, der bis auf eine Zeittransformation ein Wienerprozeß ist, und man so einen asymptotisch verteilungsfreien Test konstruieren kann. Der Kern K_θ hängt in komplizierter Weise von θ ab. Das ist der Grund dafür, dass dieses Verfahren bisher für konkrete Verteilungsfamilien nur in Haywood and Khmaladze (2008) für den relativ simplen Fall der Exponentialverteilung angewendet wurde. Ein Beleg für die Schwierigkeiten, die mit der Anwendung der Khmaladze-Transformation verbunden sind, scheint die Tatsache zu sein, dass verschiedene Autoren, die für die asymptotischen Tests notwendigen Quantile durch Bootstraptechniken approximieren. Das ist ein praktikables Verfahren, dem aber der Nachteil gegenübersteht, dass man keine Einsicht in die Güte der Tests unter lokalen Alternativen hat. Wir führen in dieser Arbeit ein prinzipiell neues Verfahren ein, mit dem die Zusatzterme, die durch das Mitschätzen des unbekannten Parameters entstehen, eliminiert werden können, wenn man den Maximum-Likelihood-Schätzer verwendet. Die Eliminierung des störenden Parameters erfolgt durch eine geeignete Randomisierung des geschätzten empirischen Prozesses. Dieser Ansatz ist neu und unterscheidet sich prinzipiell von den bisherigen Zugängen in der Literatur. Für dieses neue Testverfahren lässt sich die asymptotische Güte unter lokalen Alternativen angeben. Ein zentrales Ergebnis ist, dass die Güte von Anpassungstests nur von der Projektion der Tangente der lokalen Alternativen auf das orthogonale Komplement des Tangentialraumes des Modell abhängt. Für den Cramer-von-Mises Test werden explizite Formeln für die asymptotische Güte angegeben und die Richtung mit maximaler Güte im Raum der lokalen Alternativen kann numerisch bestimmt werden.

1 Einleitung

An dieser Stelle möchte ich mich bei allen bedanken, die zum Werden dieser Arbeit beigetragen haben. Ich danke sehr herzlich Herrn Prof. Dr. Friedrich Liese, der mich stets durch seine hilfreiche Betreuung unterstützt hat. Weiterhin möchte ich mich beim DFG-Graduiertenkolleg Welisa bedanken, dafür dass ich an diesem interessanten Forschungsprojekt teilnehmen konnte und für die finanzielle Unterstützung. Mein besonderer Dank gilt Thomas für seine Unterstützung bei programmtechnischen Fragen, für seine Geduld und sein Verständnis.

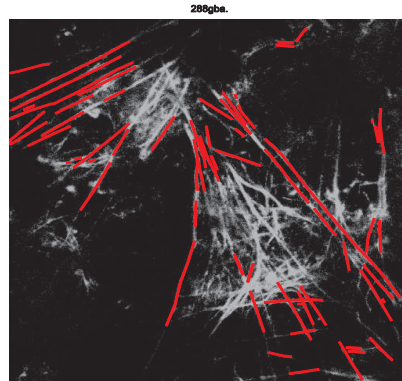
2 Modellierung von Aktinfilamenten

2.1 Der Streckenprozess der Aktinfilamente

Aktin ist ein Strukturprotein und Bestandteil des Zytoskeletts. In der Zelle bildet Aktin dynamische Filamente (die sogenannten Aktinfilamente). Diese Filamente dienen der Stabilisierung der äußeren Zellform.

Um die Zellstrukturen und das Wachstum der Zelle auf unterschiedlichen Oberflächen zu untersuchen, betrachten wir die Struktur bzw. das Verhalten der Aktinfilamente auf verschiedenen Titanoberflächen. Das folgende Bild von Aktinfilamenten wurde von Herrn Birkholz im Rahmen der Arbeit an seiner Dissertation mit Hilfe der Gratmethode gewonnen.

Abbildung 2.1: Die roten Strecken sind markierte Aktinfilamente. Verwendet wurde die Bilderkennungsmethode von Herrn Birkholz. Die Winkel zwischen den Strecken und der x-Achse wurden gemessen. Sie liegen im Intervall $[0^\circ, 180^\circ)$.



Dem Bild in der Abbildung 2.1 können wir für die weitere mathematische Modellierung folgende Aussagen entnehmen.

- Innerhalb eines Clusters bilden die Filamente zufällige Strecken.
- Die den einzelnen Filamenten entsprechenden Strecken haben eine zufällige Länge und eine zufällige Orientierung.
- Innerhalb eines Clusters gibt es eine Hauptrichtung, entlang derer die Länge der zugehörigen Filamente besonders stark ausgeprägt ist.

Für die Modellierung der zufälligen Strecken bezeichnen wir durch n die Anzahl der Strecken und durch $P_{1,j} = (X_{1,j}, Y_{1,j})$ bzw. $P_{2,j} = (X_{2,j}, Y_{2,j})$, $j = 1, \dots, n$ den Anfangspunkt bzw. den Endpunkt der jeweiligen Strecken. Weil die Strecken keine Orientierung enthalten, also der Anfangspunkt und der Endpunkt stochastisch nicht zu unterscheiden sind, setzen wir voraus

$$\mathcal{L}(X_{1,j}, Y_{1,j}, X_{2,j}, Y_{2,j}) = \mathcal{L}(X_{2,j}, Y_{2,j}, X_{1,j}, Y_{1,j}), \quad j = 1, \dots, n, \quad (2.1)$$

wobei die beiden Vektoren $(X_{1,j}, Y_{1,j})$ und $(X_{2,j}, Y_{2,j})$ nicht unbedingt unabhängig sein müssen. Für die Untersuchung der Orientierung der Strecken und den Nachweis möglicher Hauptrichtungen bilden wir die Differenzen

$$(X_j, Y_j) = (X_{2,j} - X_{1,j}, Y_{2,j} - Y_{1,j}), \quad j = 1, \dots, n. \quad (2.2)$$

2 Modellierung von Aktinfilamenten

Aus (2.1) ergibt sich

$$\mathcal{L}(X_j, Y_j) = \mathcal{L}(-X_j, -Y_j), \quad j = 1, \dots, n. \quad (2.3)$$

Die X_j und Y_j haben also symmetrische Verteilungen und insbesondere den Erwartungswert Null, sofern dieser existiert.

Wir unterscheiden jetzt zwei Typen von Modellen.

Modell 2.1 Es wird vorausgesetzt, dass (2.1) gilt und folgende Bedingung erfüllt ist

$$(X_j, Y_j), \quad j = 1, \dots, n \text{ sind u.i.v..} \quad (2.4)$$

Bei der Konstruktion von Tests für den Nachweis von Hauptrichtungen werden wir uns oft auf den folgenden Spezialfall beschränken.

Die Vektoren $(X_{i,j}, Y_{i,j})$, $i = 1, 2$, $j = 1, \dots, n$ sind unabhängig und identisch verteilt. Der Vektor $(X_{i,j}, Y_{i,j})$ besitzt den Erwartungswertvektor (μ, ν) und die Kovarianzmatrix

$$\begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}. \quad (2.5)$$

In diesem Falle gilt (2.4) und die (X_j, Y_j) besitzen eine Verteilung mit Erwartungswertvektor Null und der Kovarianzmatrix

$$\Sigma = \begin{pmatrix} \sigma^2 & \rho\sigma\tau \\ \rho\sigma\tau & \tau^2 \end{pmatrix} = 2 \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}. \quad (2.6)$$

Das Modell 2.1 entspricht Strecken aus einem homogenen Cluster, siehe Abbildung 2.2. Der Anfangspunkt $P_{1,j} = (X_{1,j}, Y_{1,j})$ bzw. der Endpunkt $P_{2,j} = (X_{2,j}, Y_{2,j})$, $j = 1, \dots, n$ besitzen die gleiche, zweidimensionale Normalverteilung mit dem Erwartungswertvektor (μ, ν) und die Kovarianzmatrix (2.5).

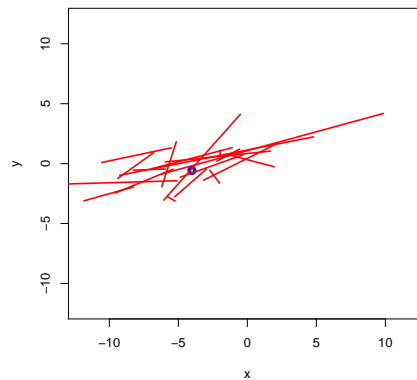


Abbildung 2.2: Die Strecken wurden aus einem homogenen Cluster durch Computersimulation erzeugt.

In einem erweiterten Modell lassen wir zu, dass die zunächst identisch verteilten Endpunkte der Strecken durch Ähnlichkeitstransformationen in der Ebene verteilt werden. Weil sich Translationen der Strecken durch die Differenzbildung in (2.2) herausheben, ergeben sich folgende Modellvoraussetzungen.

Modell 2.2 Es möge orthogonale O_j Matrizen derart geben, dass die Vektoren $O_j(X_j, Y_j)^T$ den Voraussetzungen in Modell 2.1 genügen.

In der Abbildung 2.3 sind drei Beispiele für das Modell 2.2 dargestellt. Es sei $(\tilde{X}_{i,j}, \tilde{Y}_{i,j})^T$, $i = 1, 2$ eine beliebige Rotation von $(X_{i,j}, Y_{i,j})^T$, $i = 1, 2$ im Uhrzeigersinn um den Winkel ψ , d. h.

$$\begin{pmatrix} \tilde{X}_{i,j} \\ \tilde{Y}_{i,j} \end{pmatrix} = O \begin{pmatrix} X_{i,j} \\ Y_{i,j} \end{pmatrix} \text{ mit } O = \begin{pmatrix} \cos \psi & \sin \psi \\ -\sin \psi & \cos \psi \end{pmatrix}, i = 1, 2.$$

Damit bilden die Strecken mit dem Anfangspunkt $\tilde{P}_{1,j} = (\tilde{X}_{1,j}, \tilde{Y}_{1,j})$ und dem Endpunkt $\tilde{P}_{2,j} = (\tilde{X}_{2,j}, \tilde{Y}_{2,j})$, $j = 1, \dots, n$ einen neuen Cluster. Dieser Cluster (blau) stellt in der Abbildung 2.3 eine Drehung des Clusters aus dem Modell 2.1 (rot) dar. Durch diese Drehung ändert sich die Länge der Strecke nicht.

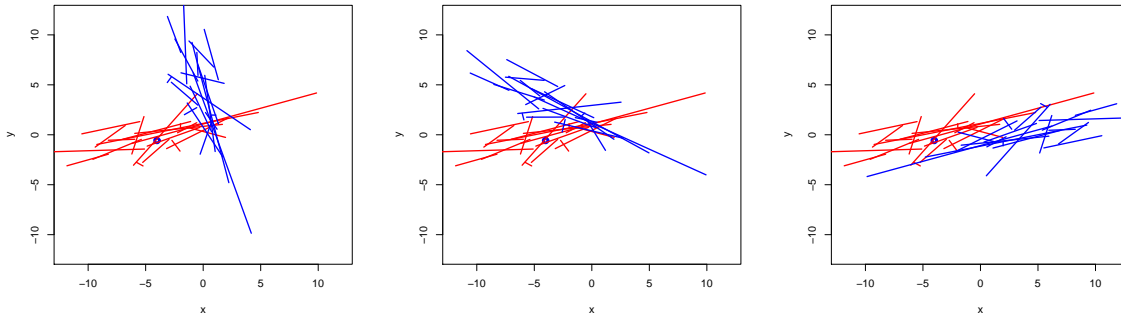


Abbildung 2.3: Der blaue Cluster ist eine Drehung des Clusters (rot) aus dem Modell 2.1 mit drei verschiedenen Winkeln $\psi = \pi/4$ (linkes Bild), $\psi = \pi/2$ (mittleres Bild) und $\psi = \pi$ (rechtes Bild).

2.2 Existenz von Hauptrichtungen

Zum Nachweis der Existenz von Richtungen, in denen die Filamente besonders stark ausgeprägt sind, verwenden wir Ideen aus dem Gebiet der linearen Regression und der Hauptkomponentenanalyse.

Sei (X, Y) ein zweidimensionaler Vektor, der endliche zweite Momente besitzt. Wir bezeichnen dann durch $\mu = EX$, $\nu = EY$ die Erwartungswerte, durch $\sigma_1^2 = V(X)$, $\sigma_2^2 = V(Y)$ die Varianzen und schließlich durch

$$\rho = \frac{\text{cov}(X, Y)}{\sigma_1 \cdot \sigma_2}$$

den Korrelationskoeffizienten. Im Problem der besten linearen Vorhersage von Y durch Beobachtung von X fragt man nach Zahlen a und b , die die Funktion

$$H(a, b) = E(Y - aX - b)^2$$

in Abhängigkeit von a und b minimieren. Diese quadratische Funktion erreicht ihr Minimum im Punkt

$$a_0 = \rho \frac{\sigma_2}{\sigma_1} \quad \text{und} \quad b_0 = \nu - \rho \frac{\sigma_2}{\sigma_1} \mu.$$

2 Modellierung von Aktinfilamenten

Die lineare Funktion

$$f(x) = \nu + \rho \frac{\sigma_2}{\sigma_1} (x - \mu)$$

heißt die lineare Regressionsfunktion. Wir betrachten jetzt das minimale mittlere Fehlerquadrat $E(Y - f(X))^2$. Unter Verwendung der Gestalt der Funktion f erhält man

$$E(Y - f(X))^2 = (1 - \rho^2) \sigma_2^2.$$

Ist ρ^2 nahe dem Wert 1, so liegt der Punkt (X, Y) nahezu auf der Regressionsgeraden $y = f(x)$.

Jetzt betrachten wir die unabhängigen Punkte $P_{1,i} = (X_{1,i}, Y_{1,i})$ bzw. $P_{2,i} = (X_{2,i}, Y_{2,i})$, deren Verteilungen mit denen von (X, Y) übereinstimmen. Fasst man jetzt $P_{1,i}$ als Anfangspunkte und $P_{2,i}$ als Endpunkte einer Strecke auf, so entstehen Strecken, die der durch die Regressionsgerade vorgegebenen Hauptrichtung tendenziell folgen.

Wir diskutieren jetzt den Einfluss der Modellparameter auf die Struktur der entstehenden Streckenprozesse, wenn die zugehörigen Strecken durch Verbinden von Paaren der Punkte $P_{1,i} = (X_{1,i}, Y_{1,i})$, $P_{2,i} = (X_{2,i}, Y_{2,i})$ entstehen. Wir setzen voraus, dass die zufälligen Punkte $(X_{i,j}, Y_{i,j})$ unabhängig und identisch verteilt sind und eine Normalverteilung mit dem Erwartungswertvektor (μ, ν) und der Kovarianzmatrix (2.5) besitzen.

Durch Betrachtung von Spezialfällen und durch Simulationen demonstrieren wir, dass die Parameter ρ , σ_1 und σ_2 die Parallelität der Strecken und die Hauptorientierung des Clusters steuern.

Zur Erzeugung von normalverteilten Pseudozufallszahlen $(X_{i,j}, Y_{i,j})$ wurde die Darstellung

$$\begin{aligned} X_{i,j} &= \sigma_1 U_{i,j} + \mu, \\ Y_{i,j} &= \nu + \rho \sigma_2 U_{i,j} + \sqrt{1 - \rho^2} \cdot \sigma_2 V_{i,j} \end{aligned}$$

verwendet, wobei die $U_{i,j}$ und $V_{i,j}$ u.i.v. und standardnormalverteilt sind.

In den unten angeführten numerischen Beispielen unterscheiden wir verschiedene Spezialfälle und wollen diskutieren, wie die Parameter ρ , σ_1 , σ_2 den Grad der Parallelität beeinflussen.

Fall 1 Für $\sigma_1 = \sigma_2$ nimmt die Parallelität der Strecken zu, falls der Korrelationskoeffizient ρ steigt. In den folgenden drei Bildern sind simulierte Strecken mit $\sigma_1 = \sigma_2 = 2$ und wachsendem ρ dargestellt. Für $\rho = 0.1$ ist die Parallelität der Strecken sehr gering, siehe linkes Bild der Abbildung 2.4. Mit steigendem ρ liegen die Strecken immer paralleler beieinander, siehe mittleres und rechtes Bild der Abbildung 2.4.

Ein Spezialfall ist $|\rho| = 1$. Dann liegen alle Strecken auf einer Geraden, siehe Abbildung 2.5. In diesem Fall sind die Strecken exakt parallel. Das liegt daran, dass alle erzeugten Punkte auf der Regressionsgeraden liegen.

Fall 2 In diesem Fall ist ρ fest. Die Parallelität der Strecken ist davon abhängig, wie stark sich σ_1^2 und σ_2^2 unterscheiden. Falls dieser Unterschied groß ist, weisen die Strecken eine hohe Parallelität auf. Das ist leicht erklärbar. Schwankt die x -Koordinate wenig und die y -Koordinate stark, so liegen die Punkte nahezu auf einer Geraden parallel zur y -Achse. Die folgenden drei Bilder verdeutlichen diese Aussage, siehe Abbildung 2.6. In diesen Beispielen sind $\rho = 0.7$ und $\sigma_2 = 1$. σ_1 hat die Werte 1, 5 bzw. 10. Mit steigendem σ_1 , nimmt die Parallelität der Strecken auf den Bildern von links nach rechts zu.

2 Modellierung von Aktinfilamenten

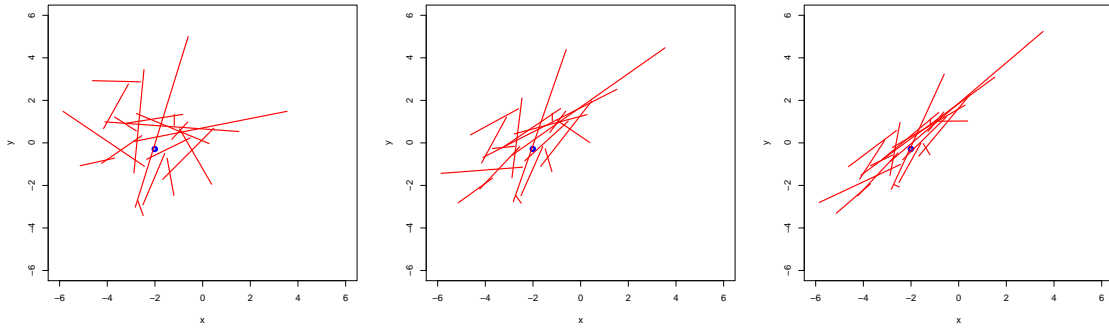


Abbildung 2.4: Simulierte Strecken: $\sigma_1 = \sigma_2 = 2$, $\rho = 0.1$ (linkes Bild), $\sigma_1 = \sigma_2 = 2$, $\rho = 0.7$ (mittleres Bild) und $\sigma_1 = \sigma_2 = 2$, $\rho = 0.9$ (rechtes Bild).

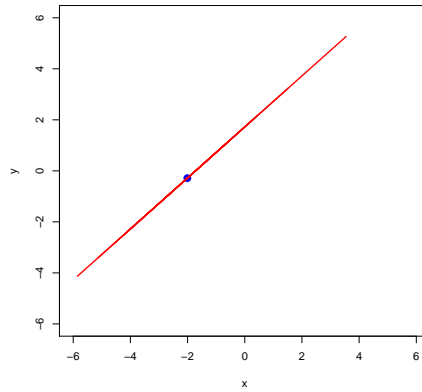


Abbildung 2.5: Simulierte Strecken mit $\sigma_2 = \sigma_1 = 2$ und $\rho = 1$.

Fall 3 Die Parameter σ_1 , σ_2 können die Hauptrichtung des Clusters beeinflussen. Je größer σ_1 , desto parallel sind die Hauptrichtung des Clusters und die x -Achse. Steigt σ_2 , dann stellt sich zunehmend eine Parallelität der Hauptrichtung der Strecken des Clusters zur y -Achse ein. In den drei Beispielen ist immer $\rho = 0.9$. Auf dem linken Bild der Abbildung 2.7 sind $\sigma_1 = 2$ und $\sigma_2 = 5$. Die Hauptrichtung des Cluster ist eher parallel zur y - als zur x -Achse. Das rechten Bild der Abbildung 2.7 zeigt den umgekehrten Fall, nämlich $\sigma_1 = 5$, $\sigma_2 = 2$. Die Hauptrichtung des Cluster dreht sich in Richtung der x -Achse.

Verwendet man zur Konstruktion der Strecken das paarweise Verbinden, so ist es wichtig, Parameter einzuführen, die ausdrücken, wie stark die entstandenen Strecken in eine bestimmte Richtung ausgerichtet sind. Die Richtung der Strecken wird durch den Vektor

$$(X_j, Y_j) = (X_{2j} - X_{1j}, Y_{2j} - Y_{1j})$$

gegeben. Durch einen Einheitsvektor $e = (\alpha, \beta)^T$ mit $\alpha^2 + \beta^2 = 1$ legen wir eine Richtung in der Ebene fest. Dann ist

$$U_{j,e} = \alpha X_j + \beta Y_j$$

die Komponente von (X_j, Y_j) in Richtung des Einheitsvektors e . Wir suchen jetzt eine Richtung e , für die das Längenquadrat der Komponente in Richtung von e im Mittel besonders groß

2 Modellierung von Aktinfilamenten

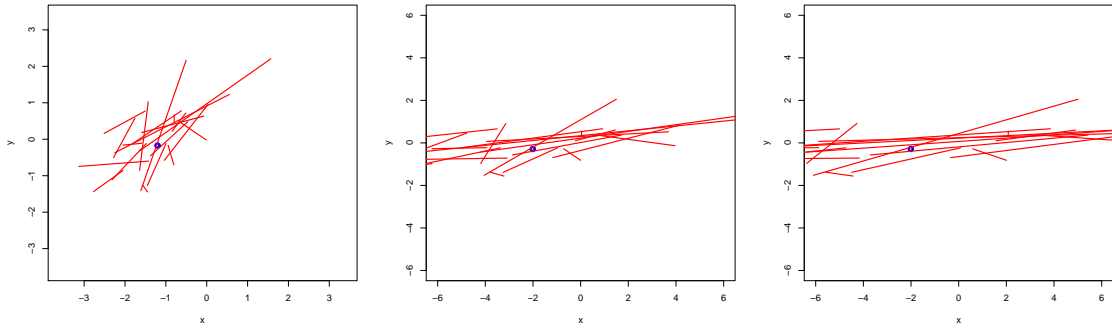


Abbildung 2.6: Simulierte Strecken: $\rho = 0.7$, $\sigma_1 = \sigma_2 = 1$ (linkes Bild), $\rho = 0.7$, $\sigma_1 = 5$, $\sigma_2 = 1$ (mittleres Bild) und $\rho = 0.7$, $\sigma_1 = 10$, $\sigma_2 = 1$ (rechtes Bild).

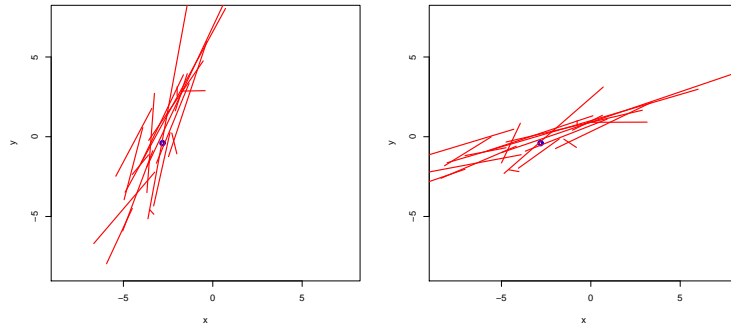


Abbildung 2.7: Simulierte Strecken: $\rho = 0.9$, $\sigma_1 = 2$, $\sigma_2 = 5$ (linkes Bild) und $\rho = 0.9$, $\sigma_1 = 5$, $\sigma_2 = 2$ (rechtes Bild).

wird, d. h. wir wollen $EU_{j,e}^2$ in Abhängigkeit von e maximieren. Wir nehmen jetzt zusätzlich an, dass die Voraussetzungen im Modell 2.1 erfüllt sind. Dann ist $EU_{j,e}^2$ unabhängig von j und kann mit Hilfe der Kovarianzmatrix Σ von (X_j, Y_j) (2.6) in folgender Form geschrieben werden

$$EU_{j,e}^2 = e^T \Sigma e.$$

Die Maximierung dieses Ausdrucks als Funktion von e führt gerade auf ein Eigenwertproblem. Es gilt

$$\max_{e: e^T e = 1} e^T \Sigma e = \lambda_1, \quad (2.7)$$

wobei λ_1 der größere der beiden Eigenwerte λ_1 , λ_2 der positiv semidefiniten und symmetrischen Matrix Σ ist, beide Eigenwerte können natürlich auch gleich sein. Bekanntlich erhält man die Eigenwerte von Σ als Lösungen der Gleichung

$$\det(\Sigma - \lambda I) = 0,$$

wobei I die Einheitsmatrix ist. In unserem Fall geht diese Gleichung in die quadratische Gleichung

$$\det \begin{pmatrix} \sigma^2 - \lambda & \rho\sigma\tau \\ \rho\sigma\tau & \tau^2 - \lambda \end{pmatrix} = (\sigma^2 - \lambda)(\tau^2 - \lambda) - (\rho\sigma\tau)^2 = 0$$

über. Die Lösungen dieser Gleichung sind

$$\begin{aligned}\lambda_1 &= \frac{1}{2}(\tau^2 + \sigma^2) + \frac{1}{2}\sqrt{(\tau^2 - \sigma^2)^2 + 4\rho^2\tau^2\sigma^2}, \\ \lambda_2 &= \frac{1}{2}(\tau^2 + \sigma^2) - \frac{1}{2}\sqrt{(\tau^2 - \sigma^2)^2 + 4\rho^2\tau^2\sigma^2}.\end{aligned}\quad (2.8)$$

Die Differenz $\lambda_1 - \lambda_2$ ist somit

$$\lambda_1 - \lambda_2 = \sqrt{(\tau^2 - \sigma^2)^2 + 4\rho^2\tau^2\sigma^2}.$$

Weiterhin gilt

$$\det \begin{pmatrix} \sigma^2 & \rho\sigma\tau \\ \rho\sigma\tau & \tau^2 \end{pmatrix} = (1 - \rho^2)\sigma^2\tau^2 = \lambda_1\lambda_2. \quad (2.9)$$

Hat man also n identisch verteilte zufällige Strecken mit den beiden Endpunkten $P_{1,j} = (X_{1,j}, Y_{1,j})$ bzw. $P_{2,i} = (X_{2,i}, Y_{2,i})$, so liegt genau dann eine dominierende Richtung für diese Strecken vor, wenn die Differenz $\lambda_1 - \lambda_2$ groß ist. Diese Aussage steht in Übereinstimmung mit den Simulationen in den Fällen 1, 2 und 3 von oben. Eine dominierende Richtung liegt bei Gleichheit der Varianzen $\sigma_1^2 = \sigma_2^2$ genau dann vor, wenn ρ^2 groß ist. Ist ρ fest gewählt, dann ist eine dominierende Richtung besonders stark ausgeprägt, wenn die Unterschiede zwischen den Varianzen σ_1^2 und σ_2^2 groß sind. Im Allgemeinen liegt bei starker Dominanz einer Richtung eine Kombination dieser beiden Fälle vor. Insgesamt können wir feststellen, dass der Nachweis einer Hauptrichtung bei zufälligen Strecken äquivalent ist zum Testproblem

$$H_0 : \lambda_1 = \lambda_2 \quad \text{gegen} \quad H_A : \lambda_1 > \lambda_2. \quad (2.10)$$

2.3 Bivariate homogene Verteilungen

Für die Analyse der zufälligen Richtung der Strecken und der Abhängigkeit zwischen der Länge der Strecken und der zugehörigen Richtung ist der Übergang zu Polarkoordinaten vorteilhaft. Für einen zufälligen Vektor (X, Y) führen wir den zufälligen Winkel Φ und die zufällige Länge R durch Übergang zu Polarkoordinaten $X = R \cos \Phi$, $Y = R \sin \Phi$ ein. Die beiden Randverteilungen von (R, Φ) geben Auskunft über die Häufigkeit der auftretenden Längen von Filamenten bzw. deren Orientierung. Dagegen gibt die gemeinsame Verteilung bzw. die bedingte Verteilung von R unter der Bedingung $\Phi = \varphi$ Auskunft darüber, in welchen Richtungen die Länge der Filamente besonders groß ist. Aus der bedingten Verteilung erhält man den bedingten Erwartungswert $E(R|\Phi = \varphi)$, der die mittlere Länge der Filamente bei gegebener Richtung φ liefert.

Die in der Literatur vorhandenen Resultate über zirkuläre Daten beziehen sich vorwiegend auf die statistische Analyse der Verteilung des zufälligen Winkels Φ . Wir verweisen auf die Bücher von Fisher et al. (1993) und Jammalamadaka und SenGupta (2001) sowie die darin zitierte, umfangreiche Literatur. Im Gegensatz zu diesen Untersuchungen wollen wir in diesem Kapitel vor allem die Wechselbeziehung zwischen der Länge R und dem Winkel Φ untersuchen. Hierzu werden wir spezielle Modelle entwickeln, innerhalb derer die Regressionsfunktion $E(R|\Phi = \varphi)$ explizit berechnet werden kann.

Wir hatten bereits untersucht, ob es eine bestimmte Richtung, beschrieben durch den Einheitsvektor $e = (e_1, e_2)$, derart gibt, dass das Quadrat des Längenanteils $(e_1 X)^2 + (e_2 Y)^2$ im Mittel besonders groß wird und nachgewiesen, dass diese Frage auf das Testproblem (2.10)

2 Modellierung von Aktinfilamenten

führt. Zur Konstruktion geeigneter Tests in finiter oder asymptotischer Form und zur Untersuchung der Regressionsfunktion $E(R|\Phi = \varphi)$, die durch Übergang zu Polarkoordinaten entsteht, sind spezielle Verteilungsannahmen notwendig. Ziel dieses Kapitels ist es, diese Verteilungsklassen einzuführen, zu diskutieren und die für die Konstruktion der Tests notwendigen Hilfsmittel bereitzustellen.

Die Polarkoordinaten in der Ebene sind definiert durch die Abbildung

$$\begin{aligned}\mathcal{M} : (0, \infty) \times [-\pi, \pi) &\rightarrow \mathbb{R} \setminus \{(0, 0)\}, \quad \text{wobei} \\ \mathcal{M}(r, \varphi) &= (r \cos \varphi, r \sin \varphi).\end{aligned}\tag{2.11}$$

Die Abbildung \mathcal{M} ist eineindeutig und durch $(\mathcal{P}_1(x, y), \mathcal{P}_2(x, y))$ bezeichnen wir die Umkehrabbildung, die wir auch in expliziter Weise mit

$$\begin{aligned}\mathcal{P}_1(x, y) &= \sqrt{x^2 + y^2}, \\ \mathcal{P}_2(x, y) &= \begin{cases} \arctan \frac{y}{x} & \text{für } x > 0 \\ \arctan \frac{y}{x} + \pi & \text{für } x < 0, y \geq 0 \\ \arctan \frac{y}{x} - \pi & \text{für } x < 0, y < 0 \\ +\pi/2 & \text{für } x = 0, y > 0 \\ -\pi/2 & \text{für } x = 0, y < 0 \end{cases}\end{aligned}\tag{2.12}$$

angeben können.

Wir verwenden im Weiteren die bekannte Transformationsregel beim Übergang zu Polarkoordinaten, siehe z. B. Hoffmann-Jørgensen (1994).

Lemma 2.3 *Ist $f : \mathbb{R}_2 \rightarrow [0, \infty]$ eine messbare Funktion, so gilt*

$$\int \left(\int f(s, t) ds \right) dt = \int_0^\infty \left(\int_0^{2\pi} f(r \cos \varphi, r \sin \varphi) r d\varphi \right) dr.\tag{2.13}$$

Sei $\mathcal{D} \subseteq \mathbb{R}_n$ eine Menge. Eine Funktion $h : \mathcal{D} \rightarrow \mathbb{R}_k$ heißt homogen der Ordnung $p \in \mathbb{R}$, falls für alle $x \in \mathcal{D}$ und alle $\lambda \geq 0$ folgt $\lambda x \in \mathcal{D}$ und

$$h(\lambda x_1, \dots, \lambda x_n) = \lambda^p h(x_1, \dots, x_n),$$

wobei wir für $p = 0$ und $\lambda = 0$ die Konvention $0^0 = 1$ verwenden. Zur Illustration der Homogenität geben wir einige einfache Beispiele homogener Funktionen von zwei Variablen

Ordnung p	$h(x, y) = a x ^p + b y ^p,$
Ordnung 1	$h(x, y) = \mathcal{P}_1(x, y) = \sqrt{x^2 + y^2},$
Ordnung 0	$h(x, y) = \mathcal{P}_2(x, y)$ in (2.12),
Ordnung $p + q$	$h(x, y) = x ^p y ^q.$

Die Abbildung 2.8 stellt den Verlauf der homogenen Funktionen

$$h(x, y) = a|x|^p + b|y|^p$$

für $a = b = 1$ und $p = 1, 2, 3$ bzw. 4 dar.

Für unsere Zwecke besonders wichtige homogene Funktionen sind der Winkel $\mathcal{P}_2(x, y)$ (0-homogen) und der Radius $\mathcal{P}_1(x, y)$ (1-homogen). Wir wollen die Struktur von homogenen Funktionen von zwei Variablen bei Übergang zu Polarkoordinaten untersuchen.

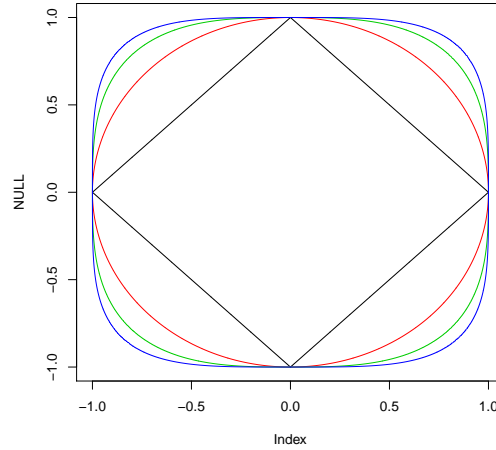


Abbildung 2.8: Verlauf der Kurve $h(x, y) = a|x|^p + b|y|^p$ mit $a = b = 1$ sowie $p = 1$ (schwarz), $p = 2$ (rot), $p = 3$ (grün) und $p = 4$ (blau).

Folgerung 2.4 Eine Funktion $h : \mathbb{R}_2 \rightarrow \mathbb{R}$ ist genau dann homogen der Ordnung p , falls mit einer gewissen Funktion $\rho : [0, 2\pi) \rightarrow \mathbb{R}$ gilt

$$h(r \cos \varphi, r \sin \varphi) = r^p \rho(\varphi). \quad (2.14)$$

Beweis. Sei h homogen der Ordnung p . Dann folgt (2.14) mit

$$\rho(\varphi) = h(\cos \varphi, \sin \varphi).$$

Umgekehrt sei für ein gegebenes p und eine Funktion $\rho : [0, 2\pi) \rightarrow \mathbb{R}$ die Funktion h definiert durch

$$h(x, y) = (\mathcal{P}_1(x, y))^p \rho(\mathcal{P}_2(x, y)).$$

Dann folgt $h(\lambda x, \lambda y) = \lambda^p h(x, y)$, weil $\mathcal{P}_2(x, y)$ 0-homogen und $\mathcal{P}_1(x, y)$ 1-homogen ist. ■

Im Hinblick auf (2.14) und zur Reduktion großer FormelAusdrücke fassen wir Funktionen $h(\cos \varphi, \sin \varphi)$ auf dem Einheitskreis direkt als Funktionen des Winkels auf und umgekehrt.

Konvention 2.5 Für Funktionen $h(\cos \varphi, \sin \varphi)$ und $k(\cos \varphi, \sin \varphi)$ auf dem Einheitskreis seien

$$h(\varphi) = h(\cos \varphi, \sin \varphi) \text{ und } k(\varphi) = k(\cos \varphi, \sin \varphi).$$

Die folgende Klasse von Verteilungen wurde für beliebige Dimension im Abschnitt 8.26 in Hoffmann-Jørgensen (1994) eingeführt und untersucht. Da wir uns vollständig auf den zweidimensionalen Fall konzentrieren, benötigen wir nicht die in Hoffmann-Jørgensen (1994) verwendeten Resultate aus der geometrischen Maßtheorie, sondern können die von uns benutzten Resultate direkt mit Hilfe von Polarkoordinaten ableiten.

Definition 2.6 Die Verteilung des zweidimensionalen zufälligen Vektors (X, Y) heißt homogen, wenn sie absolut stetig bezüglich des Lebesguemaßes ist und die Dichte die Gestalt

$$f_{X,Y}(x, y) = h_{g,h,k}(x, y) := g(h(x, y))k(x, y) \quad (2.15)$$

2 Modellierung von Aktinfilamenten

hat, wobei $h : \mathbb{R}_2 \rightarrow [0, \infty)$ mit $h(x, y) > 0$, $(x, y) \neq (0, 0)$ und $k : \mathbb{R}_2 \rightarrow [0, \infty)$ messbare homogene Funktionen der Ordnung p bzw. q sind. $g : \mathbb{R} \rightarrow [0, \infty)$ ist eine nichtnegative messbare Funktion mit

$$\int_0^\infty t^{\frac{q-p+2}{p}} g(t) dt = 1. \quad (2.16)$$

Die Verteilung mit der Dichte in (2.15) wird durch $H_{g,h,k}$ bezeichnet. g heißt die generierende Funktion der Dichte.

Bemerkung 2.7 Die in (2.15) auftretenden Funktionen g, h, k sind nicht eindeutig bestimmt. Man kann g durch Cg und k durch $C^{-1}k$ ersetzen, ohne die Dichte zu ändern. Ist die rechte Seite in (2.15) eine Dichte, so folgt aus (2.13) und der Homogenität von h und k

$$\begin{aligned} 1 &= \int_0^\infty \left(\int_0^{2\pi} g(h(r \cos \varphi, r \sin \varphi)) k(r \cos \varphi, r \sin \varphi) r dr \right) d\varphi, \\ &= \int_0^\infty \left(\int_0^{2\pi} g(r^p h(\varphi)) k(\varphi) r^{q+1} dr \right) d\varphi \quad \text{und damit} \\ &\int_0^\infty t^{\frac{q-p+2}{p}} g(t) dt < \infty \quad \text{und} \quad \int_0^{2\pi} \frac{k(\varphi)}{h^{(q+2)/p}(\varphi)} d\varphi < \infty. \end{aligned}$$

Sind umgekehrt diese beiden Bedingungen erfüllt für homogene Funktionen h, k und eine beliebige nichtnegative messbare Funktion g erfüllt, so ist $g(h(x, y))k(x, y)$ eine nichtnegative integrierbare Funktion, die zu einer Dichte normiert werden kann. Die Forderung in (2.16) ist deshalb keine Einschränkung.

Wir betrachten jetzt wichtige Spezialfälle und starten mit der homogenen Funktion $(x, y) \mapsto (|x|^p + |y|^p)$, $p > 0$. Sei g eine nichtnegative Funktion mit

$$\int_0^\infty t^{\frac{2-p}{p}} g(t) dt = 1. \quad (2.17)$$

Wir wählen jetzt die Funktion $k(x, y) = C(p)$ als Konstante so, dass gilt

$$\int \left(\int C(p) g(|x|^p + |y|^p) dx \right) dy = 1.$$

Wegen

$$\begin{aligned} 1 &= C(p) \int_0^\infty \left(\int_0^{2\pi} g(r^p (|\cos \varphi|^p + |\sin \varphi|^p)) r dr \right) d\varphi \\ &= C(p) \left(\int_0^\infty g(t) t^{\frac{2}{p}-1} dt \right) p^{\frac{2}{p}} \int_0^{2\pi} (|\cos \varphi|^p + |\sin \varphi|^p)^{-\frac{2}{p}} d\varphi \\ &= C(p) p^{\frac{2}{p}} \int_0^{2\pi} (|\cos \varphi|^p + |\sin \varphi|^p)^{-\frac{2}{p}} d\varphi \end{aligned}$$

ist $C(p)$ von g unabhängig. Zur Bestimmung von $C(p)$ können wir deshalb

$$g(t) = g_{N,p}(t) := p^{\frac{2}{p}} t^{\frac{2-p}{p}} \exp\{-t^p\}, \quad t \geq 0 \quad (2.18)$$

setzen und erhalten

$$\begin{aligned} \frac{1}{C(p)} &= \left(p^{\frac{2}{p}} \int_0^\infty t^{\frac{2-p}{p}} \exp\{-t^p\} dt \right)^2 \\ &= \left(p^{\frac{2}{p}} \cdot \frac{\Gamma\left(\frac{2-p}{p^2} + \frac{1}{p}\right)}{p} \right)^2. \end{aligned}$$

Mit diesem $C(p)$ setzen wir

$$f_{X,Y}(x, y) = h_{g,p}(x, y) = C(p)g(|x|^p + |y|^p) \quad (2.19)$$

und zur Abkürzung $H_{g,p} = H_{g,h_p,k}$. Für $p = 2$ wird die Verteilung $H_{g,2}$ als *sphärisch symmetrische Verteilung* bezeichnet. Diese Verteilungen haben also die Lebesguedichte

$$h_{g,2}(x, y) = C(2)g(x^2 + y^2). \quad (2.20)$$

Die zweidimensionale Normalverteilungsdichte $(2\pi)^{-1} \exp\{-(x^2 + y^2)/2\}$ ist ein Spezialfall, wenn man $g(t) = \frac{1}{2} \exp\{-\frac{t}{2}\}$ setzt. In Verallgemeinerung dieser speziellen, die Dichte generierenden Funktion g wird in Arbeiten über elliptisch konturierte Verteilungen oft die Klasse der Kotz-Verteilungen betrachtet. Diese dreiparametrische Klasse von Dichten verallgemeinert gleichzeitig die Gammaverteilung und die Weibullverteilung. Die Dichten sind gegeben durch

$$g_{\alpha,\beta,\gamma}(t) = C(\alpha,\beta,\gamma)t^{\alpha-1} \exp\{-\beta t^\gamma\}, \quad \alpha, \beta, \gamma > 0, t \geq 0. \quad (2.21)$$

Eine weitere wichtige Klasse von Funktionen, welche die Dichte generieren, ist die Klasse der Pearson-Typ-VII-Funktionen

$$\pi_{\alpha,\beta}(t) = C(\alpha,\beta)(1 + \beta t)^{-(1+\alpha)}, \quad \alpha \geq \frac{3}{2}, \beta > 0, t \geq 0,$$

wobei für $g_{\alpha,\beta,\gamma}(t)$ und $\pi_{\alpha,\beta}(t)$ die Konstanten so zu wählen sind, dass die Bedingung (2.16) erfüllt ist. Weil die Flanken von $\pi_{\alpha,\beta}$ nur polynomial abfallen, werden diese Funktionen zur Modellierung schwerer Flanken verwendet.

Für $p = 1$ wird $H_{g,1}$ als l_1 -Norm-symmetrische Verteilung bezeichnet, siehe Fang and Zhang (1990). Die Verteilungen $H_{g,1}$ und $H_{g,2}$ für allgemeine Dimension wurden in einer Vielzahl von Veröffentlichungen untersucht. Wir verweisen hier auf die Standardwerke Fang and Zhang (1990) und Fang et al. (1990). Die Verteilungen $H_{g,p}$ für beliebige Dimension n werden als $l_{n,p}$ -symmetrische Verteilungen bezeichnet, siehe Richter (2009). Die wahrscheinlichkeitstheoretische Struktur der Verteilungen $H_{g,p}$ wurde für beliebige Dimension ebenfalls in einer Vielzahl von Veröffentlichungen untersucht. Als eine Auswahl verweisen wir für Details auf folgende Arbeiten: Barthe et al. (2005), Rachev and Rüschendorf (1991), Richter (1991), (2007), (2009), Schechtman and Zinn (1990), Song and Gupta (1997) und Szablowski (1998).

Die im folgenden Satz formulierte Unabhängigkeit und die Verteilungsaussagen sind bereits in Hoffmann-Jørgensen (1994) zu finden. Dort werden für den Fall allgemeiner Dimension Hilfsmittel aus der geometrischen Maßtheorie verwendet. Für den einfachen Fall der Dimension zwei geben wir hier einen direkten Beweis mit Hilfe von Polarkoordinaten.

Wir bemerken, dass die im Weiteren auftretenden Funktion h und k definiert sind durch

$$h(\varphi) = h(\cos \varphi, \sin \varphi) \text{ und } k(\varphi) = k(\cos \varphi, \sin \varphi).$$

Satz 2.8 Der zufällige Vektor (X, Y) möge die Dichte in (2.15) besitzen. Dann haben die Zufallsvariablen

$$R = \mathcal{P}_1(X, Y) = \sqrt{X^2 + Y^2} \quad \text{und} \quad \Phi = \mathcal{P}_2(X, Y),$$

die gemeinsame Dichte

$$f_{R, \Phi}(r, \varphi) = I_{[0, 2\pi)}(\varphi) I_{(0, \infty)}(r) g(r^p h(\varphi)) k(\varphi) r^{q+1}. \quad (2.22)$$

Die Zufallsvariablen

$$L = R \cdot h^{1/p}(\Phi) \quad \text{und} \quad \Phi \quad (2.23)$$

sind unabhängig. Sei $\alpha(p, q) = \frac{2+q}{p}$. Die Zufallsvariable L^p besitzt die Dichte

$$f_{L^p}(t) = I_{(0, \infty)}(t) t^{\alpha(p, q)-1} g(t). \quad (2.24)$$

Die Verteilung von L^p wird mit G bezeichnet. Die Zufallsvariable Φ besitzt die Dichte

$$f_{h, k}(\varphi) = I_{[0, 2\pi)}(\varphi) \frac{k(\varphi)}{p h^{\alpha(p, q)}(\varphi)}. \quad (2.25)$$

Die zugehörige Verteilung wird mit $P_{h, k}$ bezeichnet.

Beweis. Für eine beliebige messbare Funktion $\psi : [0, \infty) \times [0, 2\pi) \rightarrow [0, \infty)$ folgt aus (2.15)

$$\begin{aligned} E\psi(R, \Phi) &= \int \left(\int \psi(\mathcal{P}_1(x, y), \mathcal{P}_2(x, y)) f_{X, Y}(x, y) dx \right) dy \\ &= \int \left(\int \psi(\mathcal{P}_1(x, y), \mathcal{P}_2(x, y)) g(h(x, y)) k(x, y) dx \right) dy. \end{aligned} \quad (2.26)$$

Mit Hilfe von (2.13) und der Homogenität von h und k ergibt sich

$$E\psi(R, \Phi) = \int \left(\int \psi(r, \varphi) g(r^p h(\varphi)) k(\varphi) r^{q+1} dr \right) d\varphi.$$

Hieraus folgt (2.22). Mit Hilfe der Variablensubstitution

$$r = \left(\frac{t}{h(\varphi)} \right)^{\frac{1}{p}} \quad \text{und} \quad r^{q+1} dr = \frac{1}{p} t^{\frac{q+2}{p}-1} h^{-\frac{q+2}{p}}(\varphi) dt$$

und der q -Homogenität von k erhalten wir für L in (2.23)

$$\begin{aligned} E\psi(L^p, \Phi) &= \int \left(\int \psi(r^p h(\varphi), \varphi) f_{R, \Phi}(r, \varphi) dr \right) d\varphi \\ &= \int \left(\int \psi(t, \varphi) I_{[0, 2\pi)}(\varphi) \frac{1}{p} t^{\alpha(p, q)-1} g(t) k(\varphi) h^{-\alpha(p, q)}(\varphi) dt \right) d\varphi. \end{aligned} \quad (2.27)$$

Setzt man speziell $\psi(L, \Phi) = 1$, so ergibt sich wegen $\int_0^\infty t^{\alpha(p, q)-1} g(t) dt = 1$

$$\left(\int_0^{2\pi} \frac{k(\varphi)}{p h^{\alpha(p, q)}(\varphi)} d\varphi \right) \left(\int_0^\infty t^{\alpha(p, q)-1} g(t) dt \right) = \int_0^{2\pi} \frac{k(\varphi)}{p h^{\alpha(p, q)}(\varphi)} d\varphi = 1.$$

Damit folgt, dass die Funktionen in (2.24) und (2.25) Wahrscheinlichkeitsdichten sind. Setzt man $\psi(L, \Phi) = I_A(L) I_B(\Phi)$ für zwei Borelmengen A und B , so folgt aus (2.27), dass die Zufallsvariablen L^p und Φ und damit auch L und Φ unabhängig sind und die angegebenen Dichten besitzen. ■

Wir berechnen jetzt bedingte Dichten und den bedingten Erwartungswert von R unter der Bedingung $\Phi = \varphi$.

Folgerung 2.9 Es seien die Voraussetzungen von Satz 2.8 erfüllt. Die Zufallsvariablen R und R^p besitzen die Dichten

$$\begin{aligned} f_R(r) &= I_{(0,\infty)}(r) r^{q+1} \int_0^{2\pi} g(r^p h(\psi)) k(\psi) d\psi, \\ f_{R^p}(t) &= I_{(0,\infty)}(t) \frac{1}{p} t^{\frac{q}{p}} \int_0^{2\pi} g(th(\psi)) k(\psi) d\psi. \end{aligned}$$

Die bedingten Dichten haben folgende Darstellungen

$$\begin{aligned} f_{R|\Phi}(r|\varphi) &= I_{(0,\infty)}(r) p g(r^p h(\varphi)) h^{\alpha(p,q)}(\varphi) r, \quad 0 \leq \varphi < 2\pi, \\ f_{\Phi|R}(\varphi|r) &= I_{[0,2\pi)}(\varphi) \frac{g(r^p h(\varphi)) k(\varphi)}{\int_0^{2\pi} g(r^p h(\psi)) k(\psi) d\psi}, \quad r > 0. \end{aligned}$$

Weiterhin gilt

$$E(R|\Phi = \varphi) = \left(\int_0^\infty g(t) t^{\frac{3}{p}-1} dt \right) h^{\frac{q-1}{p}}(\varphi).$$

Beweis. Die Zufallsvariable R besitzt wegen (2.22) die Dichte

$$f_R(r) = \int_0^{2\pi} I_{(0,\infty)}(r) g(r^p h(\psi)) k(\psi) r^{q+1} d\psi.$$

Durch Variablentransformation erhält man

$$\begin{aligned} f_{R^p}(t) &= \frac{1}{p} t^{-1/p} f_R(t^{1/p}) \\ &= I_{(0,\infty)}(t) \frac{1}{p} t^{\frac{q}{p}} \int_0^{2\pi} g(th(\psi)) k(\psi) d\psi. \end{aligned}$$

Die bedingte Dichte $f_{R|\Phi}(r|\varphi)$ ergibt sich als Quotient der Dichten in (2.22) und (2.25). Mit Hilfe der bedingte Dichte $f_{R|\Phi}(r|\varphi)$ erhalten wir

$$\begin{aligned} E(R|\Phi = \varphi) &= \int_0^\infty r f_{R|\Phi}(r|\varphi) dr \\ &= p h^{\alpha(p,q)}(\varphi) \int_0^\infty r^2 g(r^p h(\varphi)) dr. \end{aligned}$$

Die Variablensubstitution

$$r = \left(\frac{t}{h(\varphi)} \right)^{\frac{1}{p}} \quad \text{und} \quad dr = \frac{1}{p} t^{\frac{1-p}{p}} h^{-\frac{1}{p}}(\varphi) dt$$

und $\alpha = \frac{2+q}{p}$ ergeben

$$\begin{aligned} E(R|\Phi = \varphi) &= h^{\alpha(p,q)}(\varphi) \int_0^\infty \frac{t^{2/p}}{h^{2/p}(\varphi)} g(t) t^{\frac{1-p}{p}} h^{-\frac{1}{p}}(\varphi) dt \\ &= h^{\frac{q-1}{p}}(\varphi) \int_0^\infty t^{\frac{3}{p}-1} g(t) dt. \end{aligned}$$

■

2 Modellierung von Aktinfilamenten

In der obigen Darstellung des bedingten Erwartungswertes ist das eingehende Integral immer definiert, kann aber den Wert ∞ haben. Wir lassen also, wie im Rahmen nichtnegativer Zufallsvariablen üblich, zu, dass Erwartungswerte und bedingte Erwartungswerte auch Unendlich sein können.

Wir wollen jetzt eine stochastische Darstellung für einen zufälligen Vektor mit einer homogenen Verteilung geben. Aus $X = R \cos \Phi$, $Y = R \sin \Phi$ und (2.23) erhalten wir unmittelbar folgende Aussage.

Folgerung 2.10 *Besitzt der zufällige Vektor (X, Y) eine homogene Verteilung mit der Dichte (2.15), dann hat er die Darstellung*

$$\begin{pmatrix} X \\ Y \end{pmatrix} = L \cdot h^{-1/p}(\Phi) \begin{pmatrix} \cos \Phi \\ \sin \Phi \end{pmatrix},$$

wobei $L = Rh^{1/p}(\Phi)$ und Φ unabhängige Zufallsvariablen sind.

Mit Hilfe der obigen stochastischen Darstellung können wir einfache Bedingungen finden, die die Existenz von höheren gemischten Momenten sichern. Es handelt sich bei den gemischten Momenten $EX^m Y^n$ also um die Existenz von Erwartungswerten von homogenen Funktionen H des Vektors (X, Y) .

Folgerung 2.11 *$H(x, y)$ sei eine messbare und homogene Funktion der Ordnung β und $\mathcal{H}(\varphi) = H(\cos \varphi, \sin \varphi)$. Sei $\mathcal{L}(X, Y) = F_{g,h,k}$. Dann gilt*

$$E|H(X, Y)| < \infty$$

genau dann, wenn

$$EL^\beta = \int_0^\infty s^{\alpha(p,q)-1+\frac{\beta}{p}} g(s) ds < \infty \quad \text{und} \quad (2.28)$$

$$E|h^{-\beta/p}(\Phi)| |\mathcal{H}(\Phi)| = \int_0^{2\pi} \frac{|\mathcal{H}(\varphi)| k(\varphi)}{p h^{\alpha(p,q)+\beta/p}(\varphi)} d\varphi < \infty \quad (2.29)$$

erfüllt sind. In diesem Falle gilt

$$EH(X, Y) = [EL^\beta] [E(h^{-\beta/p}(\Phi) \mathcal{H}(\Phi))]. \quad (2.30)$$

Beweis. Die Unabhängigkeit von L und Φ ergibt

$$\begin{aligned} E|H(X, Y)| &= ER^\beta |H(\cos \Phi, \sin \Phi)| \\ &= (EL^\beta) (E|h^{-\beta/p}(\Phi) \mathcal{H}(\Phi)|). \end{aligned}$$

Es gilt

$$EL^k = E(L^p)^{\frac{k}{p}} = \int_0^\infty s^{\frac{k}{p}} f_{L^p}(s) ds.$$

Weil L^p die Dichte in (2.24) hat, folgt die erste Behauptung. Die Aussage (2.30) folgt durch Weglassen der Betragsstriche in der obigen Rechnung. ■

Wir betrachten die homogene Dichte

$$h_{g,p}(x, y) = C(p) g(|x|^p + |y|^p)$$

2 Modellierung von Aktinfilamenten

und untersuchen, wann j -te Momente existieren, d. h. $E(|x|^j + |y|^j) < \infty$ gilt. Weil $C(p)$ eine Konstante ist, also $q = 0$ gilt, folgt $\alpha(p, q) = \frac{2+q}{p} = \frac{2}{p}$.
Damit geht (2.28) über in

$$\begin{aligned} EL^j &= \int_0^\infty s^{\frac{2}{p}-1+\frac{j}{p}} g(s) ds \\ &= \int_0^\infty s^{\frac{2+j}{p}-1} g(s) ds < \infty. \end{aligned}$$

Die Bedingung (2.29) geht wegen der Konstanz von $k(\varphi)$ über in

$$\int_0^{2\pi} \frac{|\cos \varphi|^j + |\sin \varphi|^j}{(|\cos \varphi|^p + |\sin \varphi|^p)^{\frac{2+j}{p}}} d\varphi < \infty.$$

Diese Bedingung ist automatisch erfüllt, weil der Integrand nach oben und unten durch positive Konstanten beschränkt ist. Damit haben wir folgendes Ergebnis erhalten.

Lemma 2.12 Für die Dichte $h_{g,p}$ in (2.19) gilt

$$\begin{aligned} &\int \left(\int |x|^j + |y|^j h_{g,p}(x, y) dx \right) dy < \infty \\ \iff &\int_0^\infty s^{\frac{2+j}{p}-1} g(s) ds < \infty. \end{aligned}$$

Unterscheiden sich zwei homogene Verteilungen nur in den Funktionen g , so kann man die Erwartungswerte $EH(X, Y)$ einfach ineinander umrechnen, weil der erste Faktor in (2.30) nur von g und der zweite nur von h und k abhängt. Wir werden diesen Zusammenhang später für elliptisch konturierte Verteilungen noch genauer betrachten.

Folgerung 2.13 H sei eine homogene Funktion der Ordnung β . Sind (X_1, Y_1) und (X_2, Y_2) zufällige Vektoren mit den Verteilungen $F_{g_0, h, k}$ und $F_{g, h, k}$ und gilt $E|H(X_i, Y_i)| < \infty$, $i = 0, 1$, so folgt

$$EH(X_0, Y_0) = \frac{EL_0^\beta}{EL_1^\beta} EH(X, Y). \quad (2.31)$$

Wir sagen, dass die zweidimensionale Dichte $f(x, y)$ in *Standardform* vorliegt, wenn gilt

$$\begin{aligned} \int (\int x^2 f(x, y) dx) dy &= 1 \quad \text{und} \quad \int (\int y^2 f(x, y) dx) dy = 1, \\ \int (\int xy f(x, y) dx) dy &= 0, \\ \int (\int xf(x, y) dx) dy &= 0 \quad \text{und} \quad \int (\int yf(x, y) dx) dy = 0. \end{aligned} \quad (2.32)$$

Wir bezeichnen durch \mathfrak{M} die Menge der regulären 2×2 Matrizen A , durch \mathfrak{S} die Menge der positiv definiten symmetrischen Matrizen Σ und durch $\mathfrak{S}_0 \subseteq \mathfrak{S}$ die Menge der 2×2 Diagonalmatrizen mit gleichen Elementen in der Hauptdiagonalen. Wir betrachten jetzt die Familie der Dichten

$$f_A(x, y) := (|\det(A)|^{-1} f(A^{-1}(x, y)^T), \quad A \in \mathfrak{M}. \quad (2.33)$$

Ist die Normierungsbedingung (2.32) erfüllt und hat (X, Y) die Dichte f , so ist $\Sigma = AA^T$ gerade die Kovarianzmatrix des Vektor $A(X, Y)^T$. Sei

$$\mathfrak{G}_{p,q} = \{g : (0, \infty) \rightarrow (0, \infty) \text{ messbar und } \int_0^\infty t^{\frac{q-p+2}{p}} g(t) dt = 1\}.$$

2 Modellierung von Aktinfilamenten

Zu vorgegebenen h, k betrachten wir jetzt die folgende Familie von Dichten

$$\mathfrak{F}(h, k) = \{f_{g,h,k,A} : g \in \mathfrak{G}_{p,q}, A \in \mathfrak{M}\}. \quad (2.34)$$

Hat der Vektor (X, Y) die homogene Verteilung $F_{g,h,k}$ mit der Dichte in (2.15) und ist $A \in \mathfrak{M}$, dann hat $A(X, Y)^T$ die Dichte

$$f_{g,h,k,A}(x, y) := (|\det(A)|^{-1})g(h(A^{-1}(x, y)^T)k(A^{-1}(x, y)^T)).$$

Das ist wieder eine homogene Dichte, wenn man statt h und k die neuen homogenen Funktionen

$$\begin{aligned} h_A(x, y) &= h(A^{-1}(x, y)^T) \quad \text{und} \quad k_A(x, y) = (|\det(A)|)^{-1}k(A^{-1}(x, y)^T) \quad \text{bzw.} \\ h_A(\varphi) &= h_A(\cos \varphi, \sin \varphi) \quad \text{und} \quad k_A(\varphi) = k_A(\cos \varphi, \sin \varphi) \end{aligned}$$

verwendet.

Folgerung 2.14 Ist (X, Y) ein zufälliger Vektor mit der homogenen Dichte $f_{g,h,k}$ und $A \in \mathfrak{M}$, dann hat $A^{-1}(X, Y)^T$ die homogene Dichte $f_{g,h,k,A}$ und die stochastische Darstellung

$$A(X, Y)^T = (L \cdot h_A^{-1/p}(\Phi) \cos \Phi, L \cdot h_A^{-1/p}(\Phi) \sin \Phi)^T,$$

wobei L und Φ unabhängig sind. L hat die Verteilung G mit der Dichte in (2.24). Φ hat die Verteilung P_{h_A, k_A} mit der Dichte

$$f_{h_A, k_A}(\varphi) = I_{[0, 2\pi)}(\varphi) \frac{k_A(\varphi)}{ph_A^{\alpha(p,q)}(\varphi)}.$$

Ist speziell $A = al$, $a > 0$, so ändert sich die Verteilung des Winkels nicht, d. h.

$$P_{h_A, k_A} = P_{h, k} \quad \text{für } A = al, \quad a \neq 0. \quad (2.35)$$

Beweis. Die erste Behauptung folgt aus der Definition von k_A und h_A . Die zweite Aussage ergibt sich aus der Tatsache, dass sich bei reiner Streckung der Winkel nicht ändert, folgt aber auch aus der Rechnung

$$\begin{aligned} \frac{k_{al}(\varphi)}{ph_{al}^{\alpha(p,q)}(\varphi)} I_{[0, 2\pi)}(\varphi) &= \frac{|\det(al)|^{-1}k(a^{-1} \cos \varphi, a^{-1} \sin \varphi)}{ph^{\alpha(p,q)}(a^{-1} \cos \varphi, a^{-1} \sin \varphi)} I_{[0, 2\pi)}(\varphi) \\ &= \frac{a^{-2}a^{-q}k(\varphi)}{a^{-\frac{2+q}{p}}p h^{\alpha(p,q)}(\varphi)} I_{[0, 2\pi)}(\varphi) = \frac{k(\varphi)}{ph^{\alpha(p,q)}(\varphi)} I_{[0, 2\pi)}(\varphi). \end{aligned}$$

■

Wir bemerken, dass im Allgemeinen die Dichte $f_{g,h,k,A}$ explizit von A abhängt. Wir betrachten jetzt spezielle homogene Dichten, die nur von AA^T und damit nur von der Kovarianzmatrix abhängen, sofern die zweiten Momente existieren.

Definition 2.15 Sei \mathfrak{E} die Familie derjenigen Verteilungen $EL_{g,\Sigma}$, die absolut stetig bezüglich des Lebesgueschen Maßes sind und deren Dichten sich in der Form

$$\varphi_{g,\Sigma}(x, y) = \left(\pi \sqrt{\det(\Sigma)}\right)^{-1} g((x, y)\Sigma^{-1}(x, y)^T) \quad (2.36)$$

darstellen lassen, wobei Σ eine positiv definite 2×2 Matrix und g eine Wahrscheinlichkeitsdichte auf $(0, \infty)$ sind. Die Verteilungen $EL_{g,\Sigma}$ heißen elliptisch konturierte Verteilungen.

Bemerkung 2.16 Setzt man in (2.36) $k = \pi^{-1}$ und $h(t) = t^2$, so folgt $\varphi_{g,\Sigma} = f_{g,h,k,A}$, falls $\Sigma = AA^T$ gilt. Man erkennt also, dass jede elliptisch konturierte Dichte eine spezielle homogene Dichte ist. Bei festem g wird die Familie $\varphi_{g,\Sigma}$ aus der Standarddichte

$$\varphi_{g,I}(x, y) = \frac{1}{\pi} g(x^2 + y^2)$$

durch lineare Transformationen $A \in \mathfrak{M}$ erzeugt, wobei zwei verschiedene Matrizen A_1, A_2 mit $A_1 A_1^T = A_2 A_2^T$ zur gleichen Dichte führen. Wir diskutieren noch die Normierungsbedingungen (2.32) und setzen dazu $\Sigma = I$ in $\varphi_{g,\Sigma}(x, y)$. Dann gelten offensichtlich die Bedingungen in der zweiten und dritten Zeile von (2.32). Weiterhin gilt

$$\begin{aligned} \int \left(\int x^2 \varphi_{g,I}(x, y) dx \right) dy &= \frac{1}{\pi} \int_0^\infty \left(\int_0^{2\pi} r^2 \cos^2 \varphi g(r^2) r d\varphi \right) dr \\ &= \left(\frac{1}{\pi} \int_0^{2\pi} \cos^2 \varphi d\varphi \right) \left(\int_0^\infty r^3 g(r^2) dr \right) = \frac{1}{2} \int_0^\infty t g(t) dt. \end{aligned}$$

Setzt man $\int_0^\infty t g(t) dt = 2$ voraus, so sind die Normierungsbedingungen (2.32) erfüllt und damit ist die Matrix Σ die Kovarianzmatrix. Im Falle der Normalverteilung gilt

$$g(t) = I_{(0,\infty)}(t) \frac{1}{2} \exp\left\{-\frac{1}{2}t\right\} \quad (2.37)$$

und die Bedingung $\int_0^\infty t g(t) dt = 2$ ist automatisch erfüllt.

2.4 Bivariate elliptisch konturierte Verteilungen

Wir wollen in diesem Abschnitt weitere Resultate über elliptisch konturierte Verteilungen bereitstellen. Dazu spezialisieren wir zunächst die allgemeinen Resultate über homogene Verteilungen aus dem vorigen Abschnitt. Dabei ist es für unsere Zwecke günstig die quadratische Form

$$(x, y) \Sigma^{-1} (x, y)^T = \frac{1}{(1 - \rho^2)} \left[\left(\frac{x}{\sigma} \right)^2 - 2\rho \left(\frac{x}{\sigma} \right) \left(\frac{y}{\tau} \right) + \left(\frac{y}{\tau} \right)^2 \right] \quad (2.38)$$

und die dadurch definierte Ellipse in Polarkoordinaten aufschreiben. Für die positiv definite symmetrische Kovarianzmatrix

$$\Sigma = \begin{pmatrix} \sigma^2 & \rho\sigma\tau \\ \rho\sigma\tau & \tau^2 \end{pmatrix}$$

sei O die orthogonale Matrix der normierten Eigenvektoren als Spaltenvektoren. Dann hat die Hauptachsentransformation die Form

$$\begin{aligned} O^T \Sigma O &= \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}, \quad \Sigma = O \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} O^T, \\ O^T \Sigma^{-1} O &= \begin{pmatrix} \frac{1}{\lambda_1} & 0 \\ 0 & \frac{1}{\lambda_2} \end{pmatrix}, \quad \Sigma^{-1} = O \begin{pmatrix} \frac{1}{\lambda_1} & 0 \\ 0 & \frac{1}{\lambda_2} \end{pmatrix} O^T. \end{aligned}$$

Setzen wir $x = (x_1, x_2)^T$ und $y = (y_1, y_2)^T = O^T x$, so folgt

$$(x_1, x_2) \Sigma^{-1} (x_1, x_2)^T = x^T \Sigma^{-1} x = y O^T \Sigma^{-1} O y = \frac{y_1^2}{\lambda_1} + \frac{y_2^2}{\lambda_2}.$$

2 Modellierung von Aktinfilamenten

Wir schreiben die orthogonale Matrix O als eine Drehung um den Winkel ψ . Ist $\lambda_1 > \lambda_2$, dann sei e_1 der zu λ_1 gehörige Eigenvektor und e_2 der zu λ_2 gehörige Eigenvektor. Wir führen den Winkel ψ ein durch

$$e_1 = \begin{pmatrix} \cos \psi \\ \sin \psi \end{pmatrix} \quad \text{und damit} \quad e_2 = \begin{pmatrix} -\sin \psi \\ \cos \psi \end{pmatrix}.$$

Dann gilt

$$O = (e_1, e_2) = \begin{pmatrix} \cos \psi & -\sin \psi \\ \sin \psi & \cos \psi \end{pmatrix}. \quad (2.39)$$

Gilt $\lambda_1 = \lambda_2$, dann setzen wir $\psi = 0$. Weiterhin sei jetzt $x_1 = \cos \varphi$, $x_2 = \sin \varphi$. Dann ergibt sich

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} \cos \psi & -\sin \psi \\ \sin \psi & \cos \psi \end{pmatrix} \begin{pmatrix} \cos \varphi \\ \sin \varphi \end{pmatrix} = \begin{pmatrix} \cos(\varphi - \psi) \\ \sin(\varphi - \psi) \end{pmatrix}.$$

Damit haben wir folgendes Lemma erhalten.

Lemma 2.17 *Ist die Kovarianzmatrix Σ regulär, dann besitzt die Funktion*

$$\rho_\Sigma(\varphi) = \frac{1}{(1 - \rho^2)} \left[\frac{\cos^2 \varphi}{\sigma^2} - 2\rho \frac{\cos \varphi \sin \varphi}{\sigma \tau} + \frac{\sin^2 \varphi}{\tau^2} \right] \quad (2.40)$$

die Darstellung

$$\rho_\Sigma(\varphi) = \frac{\cos^2(\varphi - \psi)}{\lambda_1} + \frac{\sin^2(\varphi - \psi)}{\lambda_2}, \quad (2.41)$$

wobei der Winkel ψ definiert ist durch den Eigenvektor $(\cos \psi, \sin \psi)^T$, der zum Eigenwert λ_1 gehört, falls $\lambda_1 > \lambda_2$ gilt. Sonst ist $\psi = 0$.

Wir wollen jetzt die elliptisch konturierten Verteilungen in die Klasse der homogenen Verteilungen einordnen und vergleichen dazu die Dichte $\varphi_{g,\Sigma}$ mit der Dichte $f_{g,h,k}$ in (2.15). Dann gilt

$$\begin{aligned} \text{EL}_{g,\Sigma} &= \text{H}_{g,h,k} \quad \text{mit} \\ h(x, y) &= (x, y) \Sigma^{-1} (x, y)^T \quad \text{und} \quad k(x, y) = \frac{1}{\pi \sqrt{\det(\Sigma)}}, \\ h(\varphi) &= h(\cos \varphi, \sin \varphi) = \rho_\Sigma(\varphi) \quad \text{und} \quad k(\varphi) = \frac{1}{\pi \sqrt{\det(\Sigma)}}, \\ p &= 2, \quad q = 0 \quad \text{und} \quad \alpha(p, q) = \frac{2+q}{p} = 1. \end{aligned} \quad (2.42)$$

Die Zufallsvariable L in (2.23) hat im Falle elliptisch konturierter Verteilungen wegen (2.40), (2.41) und (2.42) die Darstellung

$$\begin{aligned} L^2 &= R^2 \frac{1}{(1 - \rho^2)} \left[\frac{\cos^2 \Phi}{\sigma^2} - 2\rho \frac{\cos \Phi \sin \Phi}{\sigma \tau} + \frac{\sin^2 \Phi}{\tau^2} \right] \\ &= R^2 \rho_\Sigma(\Phi) = R^2 \left[\frac{\cos^2(\Phi - \psi)}{\lambda_1} + \frac{\sin^2(\Phi - \psi)}{\lambda_2} \right], \end{aligned}$$

wobei der Winkel ψ in (2.39) definiert ist und $\lambda_1 \geq \lambda_2$ die beiden Eigenwerte der Matrix Σ sind. Weil elliptisch konturierte Verteilungen spezielle homogene Verteilungen sind, wobei h und k wie in (2.42) spezialisiert sind, ergibt sich aus Satz 2.8 folgende Aussage.

Satz 2.18 Es sei $\mathcal{L}(X, Y) \in \mathfrak{E}$ mit der Dichte $\varphi_{g, \Sigma}$ und $X = R \cos \Phi$, $Y = R \sin \Phi$. Dann hat der zufällige Vektor (R, Φ) die Lebesguedichte

$$f_{R, \Phi}(r, \varphi) = \frac{1}{\pi \sqrt{\lambda_1 \lambda_2}} l_{[0, 2\pi)}(\varphi) l_{(0, \infty)}(r) r g(r^2 \rho_{\Sigma}(\varphi)) \quad (2.43)$$

mit $\rho_{\Sigma}(\varphi)$ aus (2.40). Die Zufallsvariablen $L = R \sqrt{\rho_{\Sigma}(\Phi)}$ und Φ sind unabhängig und die Lebesguedichten sind gegeben durch

$$f_{L^2}(t) = l_{(0, \infty)}(t) g(t), \quad (2.44)$$

$$f_{\Phi}(\varphi) = \frac{1}{2\pi} l_{[0, 2\pi)}(\varphi) \frac{1}{\sqrt{\lambda_1 \lambda_2 \rho_{\Sigma}(\varphi)}}. \quad (2.45)$$

Die stochastische Darstellung eines Vektors mit einer homogenen Verteilung in Folgerung (2.10) geht jetzt in die folgende Darstellung für elliptisch konturierte Vektoren über. Gilt $\mathcal{L}(X, Y) \in \mathfrak{E}$ mit der Dichte $\varphi_{g, \Sigma}$, dann hat (X, Y) die stochastische Darstellung

$$\begin{pmatrix} X \\ Y \end{pmatrix} = L \begin{pmatrix} \cos \Phi / \sqrt{\rho_{\Sigma}(\Phi)} \\ \sin \Phi / \sqrt{\rho_{\Sigma}(\Phi)} \end{pmatrix}, \quad (2.46)$$

wobei $L = R \sqrt{\rho_{\Sigma}(\Phi)}$ und Φ unabhängig sind und die Dichten durch (2.44) und (2.45) gegeben sind.

Aus Folgerung 2.9 und (2.42) erhalten wir die bedingten Dichten.

Es sei $\mathcal{L}(X, Y) \in \mathfrak{E}$ mit der Dichte $\varphi_{g, \Sigma}$. Dann gilt für die bedingten Dichten

$$\begin{aligned} f_{R|\Phi}(r|\varphi) &= 2g(r^2 h(\varphi)) h(\varphi) r, \\ f_{\Phi|R}(\varphi|r) &= \frac{g(r^2 h(\varphi))}{\int_0^{2\pi} g(r^2 h(\psi)) d\psi}. \end{aligned}$$

Weiterhin gilt

$$E(R|\Phi = \varphi) = \left(\int_0^{\infty} t^{1/2} g(t) dt \right) h^{-\frac{1}{p}}(\varphi).$$

Wir betrachten jetzt den Spezialfall einer zweidimensionalen zentrierten Normalverteilung $N_{0, \Sigma}$ mit der Kovarianzmatrix

$$\Sigma = \begin{pmatrix} \sigma^2 & \rho\sigma\tau \\ \rho\sigma\tau & \tau^2 \end{pmatrix}$$

und der Lebesguedichte

$$\varphi_{\Sigma}(x, y) = \frac{1}{2\pi\sigma\tau\sqrt{1-\rho^2}} \cdot \exp \left\{ -\frac{1}{2(1-\rho^2)} \left[\frac{x^2}{\sigma^2} - 2\rho\frac{xy}{\sigma\tau} + \frac{y^2}{\tau^2} \right] \right\}. \quad (2.47)$$

Dann ist $N_{0, \Sigma}$ eine elliptisch konturierte Verteilung mit

$$g(t) = g_N(t) = \frac{1}{2} \exp\left\{-\frac{1}{2}t\right\}, \quad (2.48)$$

$$h(x, y) = (x, y) \Sigma^{-1} (x, y)^{\top} \text{ bzw. } h(\varphi) = \rho_{\Sigma}(\varphi). \quad (2.49)$$

Damit folgt

$$\begin{aligned} E(R|\Phi = \varphi) &= \frac{1}{\sqrt{\rho_{\Sigma}(\varphi)}} \cdot \frac{1}{2} \int_0^{\infty} t^{1/2} \exp\left\{-\frac{1}{2}t\right\} dt \\ &= \frac{\sqrt{2\pi}}{2} \cdot \left(\sqrt{\frac{\cos^2(\varphi - \psi)}{\lambda_1} + \frac{\sin^2(\varphi - \psi)}{\lambda_2}} \right)^{-1}. \end{aligned}$$

2 Modellierung von Aktinfilamenten

Falls $\lambda_1 = \lambda_2 = \lambda$ ist, ist der bedingte Erwartungswert

$$E(R|\Phi = \varphi) = \frac{\sqrt{2\pi}}{2} \cdot \left(\frac{1}{\lambda}\right)^{-\frac{1}{2}}$$

nicht mehr von φ abhängig. Die folgenden Bilder in Abbildung 2.9 zeigen den Verlauf des bedingten Erwartungswertes mit Normalverteilungsdichte. Die grüne Kurve stellt immer den Fall gleicher Eigenwerte dar. Die rote, gelbe und blaue Kurve stellt jeweils den Fall unterschiedlicher Eigenwerte dar.

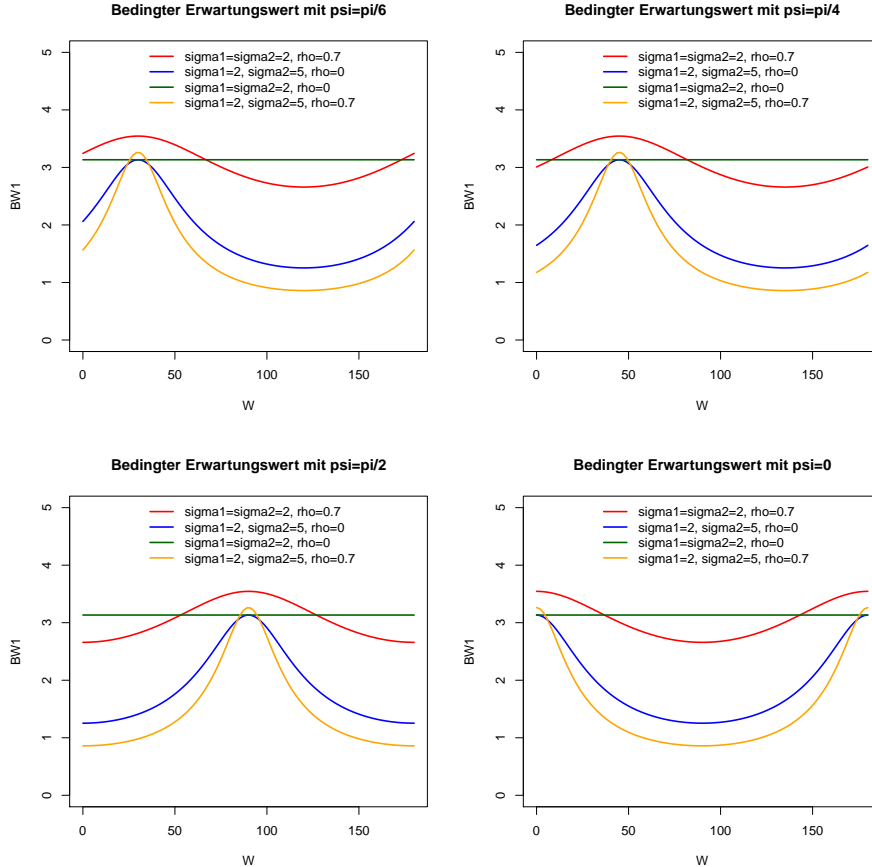


Abbildung 2.9: Verlauf des bedingten Erwartungswertes

Offensichtlich erfüllt φ_l die Normierungsbedingungen (2.32). Das erhält man natürlich auch aus der Bemerkung 2.16 wegen

$$\int_0^\infty t \frac{1}{2} \exp\left\{-\frac{1}{2}t\right\} dt = 2.$$

Wir wollen jetzt die Dichte f_{R^2} näher analysieren und erinnern, dass die Betaverteilung $B_{\alpha,\beta}$ mit den Parametern $\alpha > 0$, $\beta > 0$ die Lebesguedichte

$$b_{\alpha,\beta}(t) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} l_{(0,1)}(t) t^{\alpha-1} ((1-t)^{\beta-1}) \quad (2.50)$$

besitzt.

Lemma 2.19 Für eine beliebige messbare Funktion $h : [0, \infty) \rightarrow [0, \infty)$ und nichtnegative Zahlen a, b gilt

$$\int_0^{2\pi} h(a \cos^2 \varphi + b \sin^2 \varphi) d\varphi = 2\pi \int_0^1 h(a(1-u) + bu) b_{\frac{1}{2}, \frac{1}{2}}(u) du.$$

Beweis. Wegen

$$\sin^2 \varphi = \sin^2(\varphi + \frac{\pi}{4}) = \sin^2(\varphi + \frac{\pi}{2}) = \sin^2(\varphi + \frac{3\pi}{4})$$

gilt

$$\int_0^{2\pi} h(a \cos^2 \varphi + b \sin^2 \varphi) d\varphi = 4 \int_0^{\pi/4} h(a(1 - \sin^2 \varphi) + b \sin^2 \varphi) d\varphi.$$

Die Substitution $u = \sin^2 \varphi$ ergibt

$$d\varphi = \frac{du}{2\sqrt{u(1-u)}}$$

und mit $\Gamma(1) = 1$, $\Gamma(\frac{1}{2}) = \sqrt{\pi}$

$$\begin{aligned} \int_0^{2\pi} h(a \cos^2 \varphi + b \sin^2 \varphi) d\varphi &= 4 \int_0^1 h(a(1-u) + bu) \frac{du}{2\sqrt{u(1-u)}} \\ &= 2\pi \int_0^1 h(a(1-u) + bu) b_{\frac{1}{2}, \frac{1}{2}}(u) du. \end{aligned}$$

■

Wir untersuchen jetzt die Verteilung von R^2 .

Satz 2.20 Es sei $\mathcal{L}(X, Y) \in \mathfrak{E}$. Dann hat die Verteilung von R^2 die Dichte

$$f_{R^2}(t) = \frac{1}{\sqrt{\lambda_1 \lambda_2}} l_{(0, \infty)}(t) \int_0^1 g\left(\frac{t}{\lambda_1}(1-u) + \frac{t}{\lambda_2}u\right) b_{\frac{1}{2}, \frac{1}{2}}(u) du.$$

Folgerung 2.21 Ist $\mathcal{L}(X, Y)$ eine Normalverteilung mit Erwartungswert Null und Kovarianzmatrix Σ mit den Eigenwerten λ_1, λ_2 , so gilt

$$f_{R^2}(t) = k_{\lambda_1, \lambda_2}(t) := \frac{2}{\sqrt{\lambda_1 \lambda_2}} \int_0^1 \exp\left\{-\left(\frac{1-u}{2\lambda_1} + \frac{u}{2\lambda_2}\right)t\right\} b_{\frac{1}{2}, \frac{1}{2}}(u) du.$$

Beweis. Aus Folgerung 2.9 und (2.42) erhalten wir

$$\begin{aligned} f_{R^2}(t) &= l_{(0, \infty)}(t) \frac{1}{2\pi} \int_0^{2\pi} g(t\rho_\Sigma(\varphi)) d\varphi \\ &= f_{R^2}(t) = \frac{1}{2\pi\sqrt{\lambda_1 \lambda_2}} \int_0^{2\pi} g\left(\frac{t}{\lambda_1} \cos^2 \varphi + \frac{t}{\lambda_2} \sin^2 \varphi\right) d\varphi. \end{aligned}$$

Die Anwendung von Lemma 2.19 ergibt die Behauptung. Die Aussage in der Folgerung ergibt sich mit Hilfe der Tatsache, dass im Falle der Normalverteilung die Dichte generierende Funktion g gegeben ist durch $g(t) = 2 \exp\{-t/2\}$, siehe (2.48). ■

Das in der Definition von k_{λ_1, λ_2} auftretende Integral läßt sich nicht geschlossen berechnen. Für numerische Zwecke geben wir deshalb eine Reihenentwicklung an.

Folgerung 2.22 Die Dichte k_{λ_1, λ_2} in Folgerung 2.21 besitzt für $t > 0$ folgende Reihenentwicklungen

$$k_{\lambda_1, \lambda_2}(t) = \frac{2}{\sqrt{\lambda_1 \lambda_2}} \exp\left\{-\frac{t}{2\lambda_1}\right\} \sum_{k=0}^{\infty} \frac{t^k}{8^k (k!)^2} \frac{(2k)!}{(k-1)!} \left(\frac{\lambda_2 - \lambda_1}{\lambda_1 \lambda_2}\right)^k. \quad (2.51)$$

Beweis. Es gilt

$$\begin{aligned} \kappa_{\lambda_1, \lambda_2}(t) &= \frac{2}{\sqrt{\lambda_1 \lambda_2}} \int_0^1 \exp\left\{-\frac{t}{2\lambda_1}\right\} \times \exp\left\{\frac{ut}{2} \left(\frac{\lambda_2 - \lambda_1}{\lambda_1 \lambda_2}\right)\right\} b_{\frac{1}{2}, \frac{1}{2}}(u) du \\ &= \frac{2}{\sqrt{\lambda_1 \lambda_2}} \exp\left\{-\frac{t}{2\lambda_1}\right\} \sum_{k=0}^{\infty} \frac{t^k}{2^k k!} \left(\frac{\lambda_2 - \lambda_1}{\lambda_1 \lambda_2}\right)^k \int_0^1 u^k b_{\frac{1}{2}, \frac{1}{2}}(u) du, \end{aligned}$$

wobei das zweite Gleichheitszeichen aus dem Satz von der monotonen Konvergenz folgt. Aus der Definition der Betaverteilung erhalten wir

$$\begin{aligned} \int_0^1 u^k b_{\frac{1}{2}, \frac{1}{2}}(u) du &= \frac{\Gamma(1)}{\Gamma(\frac{1}{2})\Gamma(\frac{1}{2})} \int_0^1 u^k u^{-1/2} (1-u)^{-1/2} du \\ &= \frac{\Gamma(1)}{\Gamma(\frac{1}{2})\Gamma(\frac{1}{2})} \frac{\Gamma(k + \frac{1}{2})\Gamma(\frac{1}{2})}{\Gamma(k)}. \end{aligned}$$

Wir verwenden jetzt $\Gamma(x+1) = x\Gamma(x)$, $\Gamma(k) = (k-1)!$ und $\Gamma(\frac{1}{2}) = \sqrt{\pi}$. Dann folgt

$$\begin{aligned} \Gamma(k + \frac{1}{2}) &= (k - \frac{1}{2}) \cdots \frac{1}{2} \Gamma(\frac{1}{2}) = \frac{1}{2^k} (2k-1) \cdots 1 \sqrt{\pi} \\ &= \frac{1}{2^k} \frac{(2k)!}{k! 2^k} \sqrt{\pi} \end{aligned}$$

und

$$\int_0^1 u^k b_{\frac{1}{2}, \frac{1}{2}}(u) du = \frac{1}{k!} \frac{1}{2^k} \frac{(2k)!}{k! 2^k}.$$

Das ergibt die Reihenentwicklung. ■

Wir bemerken, dass die Reihentwicklung in Folgerung 2.22 bereits in Mathai (1982) und Sim (1992) zu finden ist.

Wir bezeichnen durch $K_{\lambda_1, \lambda_2}(t)$ die zur Dichte $k_{\lambda_1, \lambda_2}(t)$ gehörige Verteilungsfunktion. Weil $k_{\lambda_1, \lambda_2}(t)$ in $(0, \infty)$ positiv ist, ist $K_{\lambda_1, \lambda_2}(t)$ streng monoton wachsend in t . Wir bezeichnen durch $c(\gamma, \lambda_1, \lambda_2)$ das γ -Quantil von K_{λ_1, λ_2} . Wegen der Stetigkeit von K_{λ_1, λ_2} gilt also

$$K_{\lambda_1, \lambda_2}(c(\gamma, \lambda_1, \lambda_2)) = \gamma. \quad (2.52)$$

Lemma 2.23 Das Quantil $c(\gamma, \lambda_1, \lambda_2)$ ist eine stetige Funktion von (λ_1, λ_2) .

Beweis. Für positive λ_1, λ_2 und zwei Folgen $\lambda_{n,1}, \lambda_{n,2}$ gelte $\lim_{n \rightarrow \infty} \lambda_{n,i} = \lambda_i$. Sei

$$d := \liminf_{n \rightarrow \infty} c(\gamma, \lambda_{n,1}, \lambda_{n,2}) < c(\gamma, \lambda_1, \lambda_2).$$

Sei n_k eine Teilfolge mit

$$\lim_{k \rightarrow \infty} c(\gamma, \lambda_{n_k,1}, \lambda_{n_k,2}) = d.$$

Weil $K_{\lambda_1, \lambda_2}(t)$ eine stetige Funktion von $(\lambda_1, \lambda_2, t)$ ist, folgt

$$\gamma = \lim_{k \rightarrow \infty} K_{\lambda_{n_k,1}, \lambda_{n_k,2}}(c(\gamma, \lambda_{n_k,1}, \lambda_{n_k,2})) = K_{\lambda_1, \lambda_2}(d) < K_{\lambda_1, \lambda_2}(c(\gamma, \lambda_1, \lambda_2)) = \gamma.$$

2 Modellierung von Aktinfilamenten

Dieser Widerspruch ergibt

$$\liminf_{n \rightarrow \infty} c(\gamma, \lambda_{n,1}, \lambda_{n,2}) \geq c(\gamma, \lambda_1, \lambda_2).$$

Eine analoge Betrachtung ergibt

$$\limsup_{n \rightarrow \infty} c(\gamma, \lambda_{n,1}, \lambda_{n,2}) \leq c(\gamma, \lambda_1, \lambda_2),$$

was die behauptete Stetigkeit beweist. ■

3 Testen auf Hauptrichtungen des Wachstums

3.1 Testen in normalverteilten Grundgesamtheiten

3.1.1 Verteilung der empirischen Kovarianzmatrix

In diesem Kapitel wollen wir mit Hilfe statistischer Tests untersuchen, ob es Vorzugsrichtungen für die Länge gibt. Wie wir bereits im Kapitel 2 herausgearbeitet haben, läuft dies auf die Frage hinaus ob die Eigenwerte der Kovarianzmatrix gleich sind. Zu dieser klassischen Fragestellung gibt es beginnend mit Mauchly (1940a) und (1940b) zahlreiche Arbeiten, die auch den Fall einer allgemeinen Dimension erfassen. Für den Mauchly-Test und seine Modifikationen verweisen auf Anderson (1984) und die darin zitierte Literatur.

Für n u.i.v. zufällige Vektoren $(X_i, Y_i)^\top$, $i = 1, \dots, n$ mit Erwartungswert Null definieren wir die empirische Kovarianzmatrix durch

$$S_{X,Y} = \begin{pmatrix} S_{1,1} & S_{1,2} \\ S_{2,1} & S_{2,2} \end{pmatrix} = \frac{1}{n} \begin{pmatrix} \sum_{i=1}^n X_i^2 & \sum_{i=1}^n X_i Y_i \\ \sum_{i=1}^n X_i Y_i & \sum_{i=1}^n Y_i^2 \end{pmatrix}. \quad (3.1)$$

Weiterhin sei

$$R = \frac{S_{1,2}}{\sqrt{S_{1,1} S_{2,2}}}$$

der empirische Korrelationskoeffizient. Sei $N_{0,\Sigma}$ die zweidimensionale Normalverteilung mit Erwartungswert 0 und der Kovarianzmatrix

$$\Sigma = \begin{pmatrix} \sigma^2 & \rho\sigma\tau \\ \rho\sigma\tau & \tau^2 \end{pmatrix}.$$

Mit $\lambda_{2n} =$ Lebesguesches Maß im \mathbb{R}^{2n} als dominierendem Maß gilt

$$\frac{dN_{0,\Sigma}^{\otimes n}}{d\lambda_{2n}} = (2\pi \det(\Sigma))^{-n/2} \exp \left\{ -\frac{1}{2(1-\rho^2)} \left(\frac{S_{1,1}}{\sigma^2} - \frac{2\rho S_{1,2}}{\sigma\tau} + \frac{S_{2,2}}{\tau^2} \right) \right\}. \quad (3.2)$$

Damit ist $N_{0,\Sigma}^{\otimes n}$ eine dreiparametrische Exponentialfamilie und die erzeugende Statistik $(S_{1,1}, S_{1,2}, S_{2,2})$ ist nach dem Neymankriterium eine erschöpfende Statistik, siehe z. B. Satz 4.50 in Liese und Miescke (2008).

Lemma 3.1 *Gilt in (3.2) $\rho = 0$, so sind $S_{1,1}$, R und $S_{2,2}$ unabhängige Zufallsvariablen.*

Beweis. Gilt $\rho = 0$, so ist $N_{0,\Sigma}^{\otimes n}$ eine zweiparametrische Exponentialfamilie mit den natürlichen Parametern $\theta_1 = (2\sigma^2)^{-1}$, $\theta_2 = (2\tau^2)^{-1}$ und dem natürlichen Parameterbereich $\Theta = (0, \infty)^2$, der offen und nicht leer ist. Offensichtlich besitzt die Statistik R folgende Eigenschaft

$$R(ax_1, by_1, ax_2, by_2, \dots, ax_n, by_n) = R(x_1, y_1, x_2, y_2, \dots, x_n, y_n). \quad (3.3)$$

Weil für $\rho = 0$ die Verteilungen $N_{0,\Sigma}^{\otimes n}$ durch Skalentransformationen des Typs

$$(x_1, y_1, x_2, y_2, \dots, x_n, y_n) \mapsto (ax_1, by_1, ax_2, by_2, \dots, ax_n, by_n)$$

aus der Standardnormalverteilung $N_{0,I}^{\otimes n}$, $I =$ Einheitsmatrix, hervorgehen, ist R eine nichtinformative Statistik und deshalb nach den Sätzen 4.73 und 4.82 (Satz von Basu) in Liese und Miescke (2008) von der erzeugenden Statistik $(S_{1,1}, S_{2,2})$ der Exponentialfamilie

3 Testen auf Hauptrichtungen des Wachstums

unabhängig. Die Unabhängigkeit von $S_{1,1}$ und $S_{2,2}$ ist für $\rho = 0$ klar, weil die Vektoren (X_1, \dots, X_n) und (Y_1, \dots, Y_n) unabhängig sind und $S_{1,1}$ eine Funktion des ersten Vektors und $S_{2,2}$ eine Funktion des zweiten Vektors ist. ■

Wir erinnern, dass die Beta-Verteilung $B_{\alpha,\beta}$ mit den Parametern $\alpha > 0$, $\beta > 0$ durch die Lebesgue-Dichte (2.50) definiert ist. Sei

$$\gamma_\alpha(t) = I_{(0,\infty)}(t)t^{\alpha-1}\exp\{-t\}$$

die Dichte einer Gamma-Verteilung und Γ_α die zugehörige Verteilung. Der folgende Zusammenhang zwischen diesen beiden Verteilungsklassen ist bekannt und wird im Folgenden oft verwendet. Sind U_1, U_2 unabhängig und $\mathcal{L}(U_i) = \Gamma_{\alpha_i}$, $i = 1, 2$, dann gilt

$$\mathcal{L}(U_1 + U_2) = \Gamma_{\alpha_1 + \alpha_2} \quad \text{und} \quad (3.4)$$

$$\mathcal{L}\left(\frac{U_1}{U_1 + U_2}\right) = B_{\alpha_1, \alpha_2}. \quad (3.5)$$

Wir erinnern an den folgenden bekannten Zusammenhang zwischen der χ^2 -Verteilung und der Gamma-Verteilung. Sind Z_1, \dots, Z_m u.i.v. standardnormal, dann gilt

$$\mathcal{L}(Z_1^2 + \dots + Z_m^2) = \chi^2\text{-Verteilung mit } m \text{ Freiheitsgraden} = \Gamma_{m/2}. \quad (3.6)$$

Für einen linearen Unterraum $\mathbb{L} \subseteq \mathbb{R}_n$ sei $P_{\mathbb{L}}$ die Projektion auf diesen Unterraum, d. h. für jedes $x \in \mathbb{R}_n$ ist $P_{\mathbb{L}}x$ derjenige eindeutig bestimmte Vektor, der den Bedingungen

$$P_{\mathbb{L}}x \in \mathbb{L} \quad \text{und} \quad x - P_{\mathbb{L}}x \perp \mathbb{L}$$

genügt. Das nächste Lemma ist ein Spezialfall des Satzes von Cochran, siehe Anderson (1984).

Lemma 3.2 *Ist $Z = (Z_1, \dots, Z_n)$ ein standardnormalverteilter Vektor, d. h. ein normalverteilter Vektor mit Erwartungswert Null und der Einheitsmatrix als Kovarianzmatrix und $\mathbb{L} \subseteq \mathbb{R}_n$ ein k -dimensionaler linearer Unterraum, dann sind $P_{\mathbb{L}}Z$ und $Z - P_{\mathbb{L}}Z$ unabhängig, wobei $\|P_{\mathbb{L}}Z\|^2$ und $\|Z - P_{\mathbb{L}}Z\|^2$ eine χ^2 -Verteilung mit k bzw. $(n - k)$ Freiheitsgraden besitzen.*

Beweis. Es sei e_1, \dots, e_n eine orthonormale Basis von \mathbb{R}_n für die e_1, \dots, e_k eine Basis von \mathbb{L} bildet. Sei $\tilde{Z} = (e_1^T Z, \dots, e_n^T Z)$. Dann besteht \tilde{Z} aus unabhängigen standardnormalverteilten Zufallsvariablen. Wegen

$$P_{\mathbb{L}}Z = \sum_{i=1}^k (e_i^T Z)e_i \quad \text{und} \quad Z - P_{\mathbb{L}}Z = \sum_{i=k+1}^n (e_i^T Z)e_i$$

ergibt sich die Unabhängigkeit von $P_{\mathbb{L}}Z$ und $Z - P_{\mathbb{L}}Z$. Die Verteilungsaussagen ergeben sich aus (3.6). ■

Wir ermitteln jetzt die Verteilung von R^2 .

Lemma 3.3 *Gilt in (3.2) $\rho = 0$, so folgt*

$$\mathcal{L}(R^2) = B_{1/2, (n-1)/2}.$$

3 Testen auf Hauptrichtungen des Wachstums

Beweis. Wegen (3.3) können wir annehmen, dass die unabhängigen X_i, Y_i standardnormal sind. Seien $X = (X_1, \dots, X_n)$ und $Y = (Y_1, \dots, Y_n)$, h eine nichtnegative messbare Funktion und N die eindimensionale Normalverteilung. Dann gilt

$$\begin{aligned} Eh(R^2) &= Eh\left(\frac{(X^T Y)^2}{\|X\|^2 \|Y\|^2}\right) \\ &= \int \left(\int h\left(\frac{(x^T y)^2}{\|x\|^2 \|y\|^2}\right) N^{\otimes n}(dy) \right) N^{\otimes n}(dx). \end{aligned}$$

Bei festem x setzen wir $a = \frac{1}{\|x\|}x$ und berechnen das innere Integral. Es gilt

$$\int h\left(\frac{(x^T y)^2}{\|x\|^2 \|y\|^2}\right) N^{\otimes n}(dy) = Eh((a^T Y)^2 \|Y\|^{-2}). \quad (3.7)$$

Sei \mathbb{L} der lineare Unterraum, der aus den Vielfachen von a besteht. Dann gilt $P_{\mathbb{L}}Y = (a^T Y)a$. Aus dem Satz von Pythagoras folgt

$$\|Y\|^2 = \|Y - P_{\mathbb{L}}Y\|^2 + \|P_{\mathbb{L}}Y\|^2.$$

Nach Lemma 3.2 sind $U_1 = \|P_{\mathbb{L}}Y\|^2$ und $U_2 = \|Y - P_{\mathbb{L}}Y\|^2$ unabhängig und besitzen eine χ^2 -Verteilung mit einem bzw. $n - 1$ Freiheitsgraden. Die Behauptung folgt aus (3.5) und der Tatsache, dass die rechte Seite in (3.7) von a und damit von x unabhängig ist. ■

Die Aussage des folgenden Satzes einschließlich der Folgerung sind in allgemeinerer Form bekannt, siehe Anderson (1984). Für den von uns betrachteten Spezialfall gehen wir einen vereinfachten direkten Beweis.

Satz 3.4 *Gilt $\rho = 0$ in (3.2), so sind $S_{1,1}, S_{2,2}, R^2$ unabhängige Zufallsvariablen wobei $S_{1,1}/\sigma^2$ und $S_{2,2}/\tau^2$ jeweils eine χ^2 -Verteilung mit n Freiheitsgraden haben und R^2 eine Betaverteilung mit den Parametern $\alpha = 1/2$ und $\beta = (n - 1)/2$ besitzt.*

Folgerung 3.5 *Gilt $\rho = 0$ in (3.2), so haben $S_{1,1}, S_{2,2}, S_{1,2}$ die gemeinsame Dichte*

$$\begin{aligned} &f_{S_{1,1}, S_{2,2}, S_{1,2}}(s_{1,1}, s_{2,2}, s_{1,2}) \\ &= \frac{1}{4\pi(\sigma\tau)^n(n-2)!} (s_{1,1}s_{2,2} - s_{1,2}^2)^{(n-3)/2} \exp\left\{-\frac{s_{1,1}}{2\sigma^2} - \frac{s_{2,2}}{2\tau^2}\right\}, \end{aligned}$$

falls $s_{1,1} > 0, s_{2,2} > 0, s_{1,1}s_{2,2} - s_{1,2}^2 > 0$ und $f_{S_{1,1}, S_{2,2}, S_{1,2}}(s_{1,1}, s_{2,2}, s_{1,2}) = 0$ sonst.

Beweis. Der Beweis des Satzes folgt aus den Lemmata 3.1 und 3.3. Weil R eine symmetrische Verteilung hat, gilt für die Dichte f_R von R für $-1 < r < 1$

$$\begin{aligned} f_R(r) &= |r|f_{R^2}(r^2) = \frac{\Gamma(n/2)}{\Gamma((n-1)/2)\Gamma(1/2)} |r| r^{2(\frac{1}{2}-1)} (1-r^2)^{\frac{n-1}{2}-1} \\ &= \frac{\Gamma(n/2)}{\Gamma((n-1)/2)\Gamma(1/2)} (1-r^2)^{\frac{n-3}{2}}. \end{aligned}$$

Weiterhin gilt wegen (3.6)

$$\begin{aligned} f_{S_{1,1}}(s_{1,1}) &= \frac{1}{2^{\frac{n}{2}}\Gamma(\frac{n}{2})\sigma^n} s_{1,1}^{\frac{n}{2}-1} \exp\left\{-\frac{s_{1,1}}{2\sigma^2}\right\}, \\ f_{S_{2,2}}(s_{2,2}) &= \frac{1}{2^{\frac{n}{2}}\Gamma(\frac{n}{2})\tau^n} s_{2,2}^{\frac{n}{2}-1} \exp\left\{-\frac{s_{2,2}}{2\tau^2}\right\}. \end{aligned}$$

3 Testen auf Hauptrichtungen des Wachstums

Für die gemeinsame Dichte von $S_{1,1}, S_{2,2}, R$ ergibt sich aus der Unabhängigkeit dieser Zufallsvariablen

$$f_{S_{1,1}, S_{2,2}, R}(s_{1,1}, s_{2,2}, r) = D \frac{1}{(\sigma\tau)^n} (s_{1,1}s_{2,2} - s_{1,1}s_{2,2}r^2)^{(n-3)/2} \exp \left\{ -\frac{s_{1,1}}{2\sigma^2} - \frac{s_{2,2}}{2\tau^2} \right\} \sqrt{s_{1,1}s_{2,2}}$$

mit

$$D = \frac{1}{2^n \Gamma(\frac{n}{2}) \Gamma(\frac{n-1}{2}) \Gamma(\frac{1}{2})}.$$

Die Variablensubstitution $s_{1,2} = r\sqrt{s_{1,1}s_{2,2}}$ liefert bis auf die angegebene Struktur der Konstanten D die Behauptung. Die Verdoppelungsformel für die Gammafunktion

$$\Gamma(x + \frac{1}{2})\Gamma(x + 1) = \sqrt{\pi}\Gamma(2x + 1)2^{-2x} \quad (3.8)$$

ergibt mit $x + \frac{1}{2} = \frac{n-1}{2}$ die Aussage

$$\Gamma(\frac{n}{2})\Gamma(\frac{n-1}{2}) = \sqrt{\pi}\Gamma(n-1)2^{-(n-2)}.$$

Verwenden wir noch $\Gamma(\frac{1}{2}) = \sqrt{\pi}$ und $\Gamma(m) = (m-1)!$, so ergibt sich

$$D = \frac{1}{2^n \sqrt{\pi}\Gamma(n-1)2^{-(n-2)}\Gamma(\frac{1}{2})} = \frac{1}{4\pi(n-2)!}$$

und damit die Behauptung. ■

3.1.2 Invariante Teststatistiken

Sei \mathfrak{S} die Familie der positiv semidefiniten und symmetrischen 2×2 Matrizen. Weil die Statistik $(S_{1,1}, S_{1,2}, S_{2,2})$ für die Verteilungsfamilie $N_{0,\Sigma}^{\otimes n}$, $\Sigma \in \mathfrak{S}$ erschöpfend ist, können wir uns bei der Konstruktion von Teststatistiken auf Funktionen dieser erschöpfenden Statistik einschränken. Wir wollen nur solche Statistiken betrachten, die in folgendem Sinne drehungsinvariant sind.

Definition 3.6 Eine Abbildung $f : \mathfrak{S} \rightarrow \mathbb{R}_k$ heißt *drehungsinvariant*, falls für alle $A \in \mathfrak{S}$ und alle orthogonalen Matrizen O gilt

$$f(O^T A O) = f(A).$$

Weil jedes $A \in \mathfrak{S}$ durch die Hauptachsentransformation auf Diagonalgestalt gebracht wird, erhalten wir

$$f(A) = f \left(\begin{pmatrix} L_1(A) & \\ & L_2(A) \end{pmatrix} \right),$$

wobei $L_1(A) \geq L_2(A)$ die beiden Eigenwerte von A sind.

Eine Abbildung $f : \mathfrak{S} \rightarrow \mathbb{R}_k$ ist genau dann *drehungsinvariant*, wenn f eine Funktion der Eigenwerte ist. Seien $(X_{i,1}, Y_{i,2})$, $i = 1, \dots, n$ u.i.v. Vektoren mit der Verteilung $N_{0,\Sigma}$ und sei O die orthogonale Matrix mit

$$O^T \Sigma O = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} = \Lambda, \quad \lambda_1 \geq \lambda_2.$$

Wir setzen

$$\begin{pmatrix} U_i \\ V_i \end{pmatrix} = O^T \begin{pmatrix} X_i \\ Y_i \end{pmatrix}.$$

Dann sind die $(U_{i,1}, V_{i,2})$, $i = 1, \dots, n$ u.i.v. Vektoren mit der Verteilung $N_{0,\Lambda}$. Wir untersuchen jetzt den Zusammenhang der beiden empirischen Kovarianzmatrizen, die wir durch $S_{X,Y}$ bzw. $S_{U,V}$ bezeichnen. Es gilt

$$\begin{aligned} S_{U,V} &= \frac{1}{n} \sum_{i=1}^n (U_i, V_i)^T (U_{i,1}, V_{i,2}) \\ &= \frac{1}{n} \sum_{i=1}^n O^T (X_i, Y_i)^T (X_i, Y_i) O = O^T S_{X,Y} O. \end{aligned}$$

Ist jetzt $f : \mathfrak{S} \rightarrow \mathbb{R}_k$ eine drehungsinvariante Statistik, dann gilt

$$f(S_{X,Y}) = f(S_{U,V}). \quad (3.9)$$

Mit Hilfe dieser Beziehung können wir für drehungsinvariante Statistiken die Verteilung von $f(S_{X,Y})$ mit Hilfe der Verteilung von $f(S_{U,V})$ ermitteln, wobei sich die Elemente der Matrix $S_{U,V}$ mit Hilfe von unabhängigen Zufallsvariablen mit bekannten Verteilungen ausdrücken lassen.

Satz 3.7 (X_i, Y_i) , $i = 1, \dots, n$ seien u.i.v. mit Verteilung $N_{0,\Sigma}$. Es sei $f : \mathfrak{S}_2 \rightarrow \mathbb{R}_k$ eine drehungsinvariante Statistik. Dann gilt

$$\mathcal{L}(f(S_{X,Y})) = \mathcal{L}(f(S_{U,V})),$$

wobei (U, V) ein normalverteilter Vektor mit Erwartungswert Null und unabhängigen Komponenten ist, die die Varianzen λ_1, λ_2 haben, wobei $\lambda_1 \geq \lambda_2$ die Eigenwerte der Kovarianzmatrix Σ sind.

Beweis. Die Behauptung folgt aus (3.9) und der Tatsache, dass die $(U_{i,1}, V_{i,2})$, $i = 1, \dots, n$ u.i.v. Vektoren mit der Verteilung $N_{0,\Lambda}$ sind. ■

3.1.3 Likelihoodquotiententest in normalverteilten Grundgesamtheiten

Zur Konstruktion des Likelihoodquotiententests fassen wir die Dichte in (3.2) als Funktion der Kovarianzmatrix auf und bezeichnen wie üblich diese Funktion als Likelihoodfunktion, also

$$L(\Sigma) = \left(2\pi\sigma\tau\sqrt{1-\rho^2}\right)^{-n} \exp \left\{ -\frac{1}{2} \left(\frac{S_{1,1}}{\sigma^2} - 2\rho \frac{S_{1,2}}{\sigma\tau} + \frac{S_{2,2}}{\tau^2} \right) \right\}.$$

Sei

$$\mathfrak{S}_0 = \{\Sigma : \Sigma \in \mathfrak{S}, \Sigma = \sigma^2 I, \sigma^2 > 0\}.$$

Ob die unbekannte Kovarianzmatrix Σ zu einer der Teilmengen $\mathfrak{S}_0 \subseteq \mathfrak{S}_2$ gehört, lässt sich mit Hilfe von Likelihoodquotiententests testen. Die Maximierung der Likelihoodfunktion über \mathfrak{S}_0 und \mathfrak{S} lässt sich in geschlossener Form ausführen. Für den Beweis der folgenden Resultate verweisen wir auf Anderson (1984).

Lemma 3.8 Für das Testproblem

$$H_0 : N_{0,\Sigma}^{\otimes n}, \Sigma \in \mathfrak{S}_0 \quad \text{gegen} \quad H_A : N_{0,\Sigma}^{\otimes n}, \Sigma \in \mathfrak{S} \setminus \mathfrak{S}_0 \quad (3.10)$$

lautet die Likelihoodquotientenstatistik

$$\frac{\sup_{\Sigma \in \mathfrak{S}} L(\Sigma)}{\sup_{\Sigma \in \mathfrak{S}_0} L(\Sigma)} = \left(\frac{\frac{1}{2}(S_{1,1} + S_{2,2})}{\sqrt{S_{1,1}S_{2,2} - S_{1,2}^2}} \right)^n.$$

3 Testen auf Hauptrichtungen des Wachstums

Wir schreiben diese Teststatistik noch etwas um. Es gilt

$$\frac{\frac{1}{2}(S_{1,1} + S_{2,2})}{\sqrt{S_{1,1}S_{2,2} - S_{1,2}^2}} = \frac{\frac{1}{2}(S_{1,1} + S_{2,2})}{\sqrt{S_{11}S_{22}}} \frac{1}{\sqrt{1 - R^2}}. \quad (3.11)$$

Wegen $\sqrt{S_{1,1}S_{2,2}} \leq \frac{1}{2}(S_{1,1} + S_{2,2})$ mit Gleichheit für $S_{1,1} = S_{2,2}$, testet der erste Faktor in (3.11) eine mögliche Ungleichheit der Varianzen, während $1 - R^2$ eine mögliche Korrelation testet.

Für das Testproblem (3.10) wollen wir die Teststatistik mit den Eigenwerten der empirischen Kovarianzmatrix $S_{X,Y}$ in Verbindung bringen, die wir mit L_1 und L_2 bezeichnen. Dann gilt

$$L_{1(2)} = \frac{1}{2}(S_{1,1} + S_{2,2}) + (-) \sqrt{\frac{1}{4}(S_{1,1} - S_{2,2})^2 + R^2 S_{1,1} S_{2,2}}. \quad (3.12)$$

Weiterhin gilt

$$\begin{aligned} L_1 + L_2 &= S_{1,1} + S_{1,2} = \text{tr}(S_{X,Y}), \\ L_1 L_2 &= \det(S) = S_{1,1} S_{2,2} - S_{1,2}^2. \end{aligned}$$

Das ergibt

$$\frac{\sup_{\Sigma \in \mathfrak{G}} L(\Sigma)}{\sup_{\Sigma \in \mathfrak{G}_0} L(\Sigma)} = \left(\frac{\sqrt{L_1 L_2}}{\frac{1}{2}(L_1 + L_2)} \right)^{-n}. \quad (3.13)$$

Beim Likelihoodquotiententest wird die Nullhypothese in (3.10) für große Werte der linken Seite in (3.13) abgelehnt. Der Darstellung (3.13) entnimmt man, dass man äquivalent die Nullhypothese in Testproblem (3.10) auch für kleine Werte von $\sqrt{L_1 L_2} / (\frac{1}{2}(L_1 + L_2))$ ablehnen kann.

Die Verteilung der Teststatistik in (3.13) unter der Alternative wurde bereits in Girshick (1941) untersucht. Der dort gegebene Beweis ist an einigen Stellen stark verkürzt und enthält Lücken. Deshalb soll nachfolgend ein ausführlicher Beweis gegeben werden, wobei wir aber wesentliche Teile des Girshickschen Beweises verwenden. Wir fassen die positiv definiten symmetrischen Matrizen als Teilmengen von \mathbb{R}_3 auf. Jede solche Matrix Σ können wir schreiben als

$$\Sigma = \begin{pmatrix} s_{1,1} & s_{1,2} \\ s_{1,2} & s_{2,2} \end{pmatrix},$$

wobei gilt $s_{i,i} > 0$, $-\sqrt{s_{1,1}s_{2,2}} < s_{1,2} < \sqrt{s_{1,1}s_{2,2}}$ oder äquivalent $0 < s_{1,2}^2 < s_{1,1}s_{2,2}$. Die Eigenwerte sind dann gegeben durch

$$l_{1(2)} = \frac{1}{2}(s_{1,1} + s_{2,2}) + (-) \frac{1}{2} \sqrt{(s_{1,1} - s_{2,2})^2 + 4s_{1,2}^2}.$$

Sei

$$\mathcal{G} = \{(s_{1,1}, s_{2,2}, s_{1,2}) : s_{1,2}^2 < s_{1,1}s_{2,2}, 0 < s_{1,1} < \infty, 0 < s_{2,2} < \infty\}.$$

Wir betrachten die Abbildung

$$(h_1, h_2, s_{1,2}) = S(s_{1,1}, s_{2,2}, s_{1,2}). \quad (3.14)$$

Diese Abbildung bildet offensichtlich unterschiedliche Punkte auf den gleichen Bildpunkt des Bildbereichs

$$\mathcal{H} = \{(h_1, h_2, s_{1,2}) : 0 < h_2 < h_1 < \infty, 4s_{1,2}^2 < (h_1 - h_2)^2\}$$

3 Testen auf Hauptrichtungen des Wachstums

ab. Wir spalten deshalb den Bereich \mathcal{G} auf in

$$\begin{aligned}\mathcal{G}_1 &= \{(s_{1,1}, s_{2,2}, s_{1,2}) : s_{1,2}^2 < s_{1,1}s_{2,2}, 0 < s_{2,2} < s_{1,1} < \infty\}, \\ \mathcal{G}_2 &= \{(s_{1,1}, s_{2,2}, s_{1,2}) : s_{1,2}^2 < s_{1,1}s_{2,2}, 0 < s_{1,1} < s_{2,2} < \infty\}\end{aligned}$$

und bezeichnen durch S_1 und S_2 die Einschränkungen von (3.14) auf die obigen zwei Teilmengen von \mathcal{G} . Es gilt

$$\begin{aligned}s_{1,1} + s_{2,2} &= l_1 + l_2, \\ (l_1 - l_2)^2 - 4s_{1,2}^2 &= (s_{1,1} - s_{2,2})^2.\end{aligned}$$

Wir setzen

$$A = (l_1 - l_2)^2 - 4s_{1,2}^2.$$

Dann folgt, dass S_1 und S_2 eindeutige Abbildungen mit den Umkehrabbildungen

$$\begin{aligned}(s_{1,1}, s_{2,2}, s_{1,2}) &= T_1(l_1, l_2, s_{1,2}) = \frac{1}{2}((l_1 + l_2) + \sqrt{A}, (l_1 + l_2) - \sqrt{A}, 2s_{1,2}), \\ (s_{1,1}, s_{2,2}, s_{1,2}) &= T_2(l_1, l_2, s_{1,2}) = \frac{1}{2}((l_1 + l_2) - \frac{1}{2}\sqrt{A}, (l_1 + l_2) + \sqrt{A}, 2s_{1,2})\end{aligned}$$

sind. Wir berechnen die Determinante der Jacobimatrix der Abbildung T_1

$$J_{T_1}(l_1, l_2, s_{1,2}) = \frac{1}{2} \begin{pmatrix} 1 + A^{-1/2}(l_1 - l_2) & 1 - A^{-1/2}(l_1 - l_2) & 0 \\ 1 - A^{-1/2}(l_1 - l_2) & 1 + A^{-1/2}(l_1 - l_2) & 0 \\ 0 & 0 & 4 \end{pmatrix}.$$

Dann folgt zusammen mit einer entsprechenden Rechnung für T_2

$$\det(J_{T_1}(l_1, l_2, s_{1,2})) = \det(J_{T_2}(l_1, l_2, s_{1,2})) = A^{-1/2}(l_1 - l_2).$$

Wegen Satz 3.7 können wir uns bei der Berechnung der Verteilung der Eigenwerte gleich auf den Fall beschränken, wo die Kovarianzmatrix diagonal ist mit den Einträgen λ_1, λ_2 . Wir nehmen jetzt an, dass $(S_{1,1}, S_{2,2}, S_{1,2})$ die Dichte in Folgerung 3.5 hat, mit $\sigma^2 = \lambda_1, \tau^2 = \lambda_2$. Wir setzen

$$\begin{aligned}\omega_1 &= \frac{\lambda_1 + \lambda_2}{4\lambda_1\lambda_2} \quad \text{und} \quad \omega_2 = \frac{\lambda_1 - \lambda_2}{4\lambda_1\lambda_2}, \\ g(l_1, l_2) &= \frac{1}{4\pi(\lambda_1\lambda_2)^{\frac{n}{2}}(n-2)!} (l_1 l_2)^{(n-3)/2} \exp\{-\omega_1(l_1 + l_2)\}.\end{aligned}\tag{3.15}$$

Wir schreiben für $\sigma^2 = \lambda_1, \tau^2 = \lambda_2$ und $\rho = 0$ die Dichte in Satz 3.5 als Funktion von $(l_1, l_2, s_{1,2})$. Dann gilt

$$\begin{aligned}f_{S_{1,1}, S_{2,2}, S_{1,2}}(s_{1,1}, s_{2,2}, s_{1,2}) &= g(l_1, l_2) \exp\{-\omega_2\sqrt{A}\} \quad \text{auf } \mathcal{G}_1, \\ f_{S_{1,1}, S_{2,2}, S_{1,2}}(s_{1,1}, s_{2,2}, s_{1,2}) &= g(l_1, l_2) \exp\{\omega_2\sqrt{A}\} \quad \text{auf } \mathcal{G}_2.\end{aligned}$$

Sei h eine beliebige nichtnegative messbare Funktion. Dann gilt

$$Eh(L_1, L_2) = El_{\mathcal{G}_1}(S_{1,1}, S_{2,2}, S_{1,2})h(L_1, L_2) + El_{\mathcal{G}_2}(S_{1,1}, S_{2,2}, S_{1,2})h(L_1, L_2).$$

3 Testen auf Hauptrichtungen des Wachstums

Nach der Transformationsformel für Lebesgueintegrale gilt für L_1, L_2 in (3.12)

$$\begin{aligned} & Eh(L_1, L_2) \\ &= \int \left(\int \left(\int l_{\mathcal{H}}(l_1, l_2, s_{1,2}) h(l_1, l_2) g(l_1, l_2) (l_1 - l_2) A^{-1/2} \right. \right. \\ &\quad \left. \left. \times (\exp \{-\omega_1 \sqrt{A}\} + \exp \{\omega_2 \sqrt{A}\}) dl_1 \right) dl_2 ds_{1,2}. \right. \end{aligned}$$

Wir berechnen das Integral

$$\begin{aligned} J &= \int l_{\mathcal{H}}(l_1, l_2, s_{1,2}) A^{-1/2} \left(\exp \{-\omega_2 \sqrt{A}\} + \exp \{\omega_2 \sqrt{A}\} \right) ds_{1,2} \\ &= \int_{-\frac{1}{2}(l_1-l_2)}^{\frac{1}{2}(l_1-l_2)} \frac{1}{\sqrt{(l_1-l_2)^2 - 4s_{1,2}^2}} \\ &\quad \times \left(\exp \left\{ -\omega_2 \sqrt{(l_1-l_2)^2 - 4s_{1,2}^2} \right\} + \exp \left\{ \omega_2 \sqrt{(l_1-l_2)^2 - 4s_{1,2}^2} \right\} \right) ds_{1,2} \end{aligned}$$

und setzen

$$s_{1,2} = \frac{u}{2}(l_1 - l_2).$$

Dann folgt

$$\begin{aligned} J &= \int_{-1}^1 \left(\exp \left\{ -\omega_2(l_1 - l_2) \sqrt{1 - u^2} \right\} + \exp \left\{ \omega_2(l_1 - l_2) \sqrt{1 - u^2} \right\} \right) \\ &\quad \times \frac{1}{(l_1 - l_2) \sqrt{1 - u^2}} \frac{1}{2} (l_1 - l_2) du \\ &= \int_0^1 \left(\exp \left\{ -\omega_2(l_1 - l_2) \sqrt{1 - u^2} \right\} + \exp \left\{ \omega_2(l_1 - l_2) \sqrt{1 - u^2} \right\} \right) (1 - u^2)^{-1/2} du \\ &= \sum_{k=1}^{\infty} \frac{1}{(2k)!} (\omega_2(l_1 - l_2))^{2k} \int_0^1 (1 - u^2)^{k-1/2} du \\ &= \sum_{k=1}^{\infty} \frac{1}{(2k)!} (\omega_2(l_1 - l_2))^{2k} \int_0^1 (1 - t)^{k-1/2} t^{-1/2} dt, \end{aligned}$$

wobei wir für $\omega_2 = 0$ die Konvention $0^0 = 1$ verwenden. Wir benutzen jetzt die folgenden bekannten Beziehungen für die Gammaverteilung

$$\begin{aligned} \int_0^1 t^{\alpha-1} (1-t)^{\beta-1} dt &= \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}, \quad \Gamma\left(\frac{1}{2}\right) = \sqrt{\pi} \\ \Gamma\left(k + \frac{1}{2}\right) &= \frac{1}{2^k} \frac{(2k)!}{k! 2^k} \sqrt{\pi}, \end{aligned}$$

wobei sich die letzte Beziehung z. B. dadurch ergibt, dass man für eine standardnormalverteilte Zufallsvariable U einerseits die bekannte Formel für die geraden Momente verwendet und andererseits die Momente durch Rückführung auf die Gammafunktion berechnet. Insgesamt erhalten wir

$$\begin{aligned} J &= \sum_{k=0}^{\infty} \frac{1}{(2k)!} (\omega_2(l_1 - l_2))^{2k} \left(\frac{1}{2^k} \frac{(2k)!}{k! 2^k} \sqrt{\pi} \right) \sqrt{\pi} \frac{1}{k!} \\ &= \pi \sum_{k=0}^{\infty} (\omega_2(l_1 - l_2))^{2k} \frac{1}{4^k} \frac{1}{(k!)^2}. \end{aligned}$$

Damit folgt

$$\begin{aligned}
 & Eh(L_1, L_2) \\
 &= \int \left(\int l_{\mathcal{H}}(l_1, l_2, s_{1,2}) h(l_1, l_2) (l_1 - l_2) g(l_1, l_2) \pi \sum_{k=0}^{\infty} (\omega_2(l_1 - l_2))^{2k} \frac{1}{4^k} \frac{1}{(k!)^2} dl_1 \right) dl_2 \\
 &= \int \left(\int l_{\mathcal{H}}(l_1, l_2, s_{1,2}) h(l_1, l_2) (l_1 - l_2) \frac{1}{4(\lambda_1 \lambda_2)^{\frac{n}{2}} (n-2)!} (l_1 l_2)^{(n-3)/2} \right. \\
 &\quad \times \exp \{-\omega_1(l_1 + l_2)\} \sum_{k=0}^{\infty} (\omega_2(l_1 - l_2))^{2k} \frac{1}{4^k} \frac{1}{(k!)^2} dl_1 \Big) dl_2.
 \end{aligned}$$

und wir haben folgendes Resultat erhalten.

Satz 3.9 Die zufällige Vektoren $(X_i, Y_i)^T$, $i = 1, \dots, n$ seien u.i.v. mit Verteilung $N_{0, \Sigma}$. Es seien $\lambda_1 \geq \lambda_2 > 0$ die Eigenwerte von Σ . Sind L_1, L_2 durch (3.12) definiert, so hat der zufällige Vektor L_1, L_2 die Lebesguedichte

$$\begin{aligned}
 f_{L_1, L_2}(l_1, l_2) &= l_{(0, l_1)}(l_2) \frac{l_1 - l_2}{4(\lambda_1 \lambda_2)^{\frac{n}{2}} (n-2)!} (l_1 l_2)^{(n-3)/2} \exp \left\{ -\frac{\lambda_1 + \lambda_2}{4\lambda_1 \lambda_2} (l_1 + l_2) \right\} \\
 &\quad \times \sum_{k=0}^{\infty} \left(\frac{\lambda_1 - \lambda_2}{\lambda_1 \lambda_2} \right)^k \left(\frac{l_1 - l_2}{4} \right)^{2k} \frac{1}{(k!)^2}.
 \end{aligned}$$

Bemerkung 3.10 Das obige Ergebnis ist bereits in Girshick (1941) zu finden. Für die Verteilung der Eigenwerte der zufälligen Kovarianzmatrix normalverteilter Vektoren beliebiger Dimension gibt es unterschiedliche Zugänge, siehe Anderson (1984), Farrell (1985), Muirhead (1982), wobei das Haarsche Maß der orthogonalen Gruppe ein wesentliches Hilfsmittel ist. Sind die Eigenwerte der Kovarianzmatrix nicht alle gleich, dann tritt in der Dichteformel ein Integral über die orthogonale Gruppe auf. Dieses Integral lässt sich nicht geschlossen berechnen.

Wir berechnen jetzt die Dichte der Teststatistik

$$L = \frac{\sqrt{L_1 L_2}}{\frac{1}{2}(L_1 + L_2)}. \quad (3.16)$$

Die folgende Aussage wurde in Girshick (1941) bewiesen.

Satz 3.11 Die zufälligen Vektoren $(X_i, Y_i)^T$, $i = 1, \dots, n$ seien u.i.v. mit Verteilung $N_{0, \Sigma}$. Es seien $\lambda_1 \geq \lambda_2 > 0$ die Eigenwerte von Σ . Sind L_1, L_2 durch (3.12) und ist L durch (3.16) definiert, so hat L die Lebesguedichte

$$\begin{aligned}
 f_L(l) &= l_{(0,1)}(l) \frac{1}{(n-2)!} \lambda^n l^{n-2} \sum_{k=0}^{\infty} \frac{(2k+n-1)!}{(k!)^2 2^{2k}} (1-\lambda^2)^k (1-l^2)^k \\
 \text{mit } \lambda &= \frac{\sqrt{\lambda_1 \lambda_2}}{\frac{1}{2}(\lambda_1 + \lambda_2)}.
 \end{aligned}$$

Folgerung 3.12 Sind die Eigenwerte λ_1, λ_2 gleich, so besitzt L die Lebesguedichte

$$f_L(l) = l_{(0,1)}(l) (n-1) l^{n-2}. \quad (3.17)$$

Beweis. Seien

$$l = \frac{\sqrt{l_1 l_2}}{\frac{1}{2}(l_1 + l_2)} \quad \text{und} \quad \kappa = \frac{1}{2}(l_1 + l_2).$$

Dann gilt

$$\begin{aligned} \frac{l_1 - l_2}{l_1 + l_2} &= \sqrt{1 - l^2}, \\ l_1 &= \kappa(1 + \sqrt{1 - l^2}) \quad \text{und} \quad l_2 = \kappa(1 - \sqrt{1 - l^2}). \end{aligned} \quad (3.18)$$

Damit ist $S : (l_1, l_2) \mapsto (l, \kappa)$ eine eindeutige Abbildung von $\{(l_1, l_2) : 0 < l_2 < l_1 < \infty\}$ auf $(0, 1) \times (0, \infty)$ mit der Umkehrabbildung T , die durch (3.18) definiert ist. Wir berechnen die Funktionaldeterminante

$$\begin{pmatrix} \frac{\partial l_1}{\partial \kappa} & \frac{\partial l_2}{\partial \kappa} \\ \frac{\partial l_1}{\partial l} & \frac{\partial l_2}{\partial l} \end{pmatrix} = \begin{pmatrix} 1 + \sqrt{1 - l^2} & 1 - \sqrt{1 - l^2} \\ -\kappa(1 - l^2)^{-1/2}l & \kappa(1 - l^2)^{-1/2}l \end{pmatrix}.$$

Die Determinante dieser Matrix ist

$$J_T = 2\kappa(1 - l^2)^{-1/2}l.$$

Sei $K = \frac{1}{2}(L_1 + L_2)$, $\omega_1 = \frac{\lambda_1 + \lambda_2}{4\lambda_1\lambda_2}$ und $\omega_2 = \frac{\lambda_1 - \lambda_2}{4\lambda_1\lambda_2}$. Dann ist die gemeinsame Dichte von (L, K) gegeben durch

$$\begin{aligned} f_{L,K}(l, \kappa) &= \frac{2\kappa\sqrt{1 - l^2}}{4(\lambda_1\lambda_2)^{\frac{n}{2}}(n-2)!} (\kappa^2 l^2)^{(n-3)/2} 2\kappa(1 - l^2)^{-1/2}l \exp\{-\omega_1 2\kappa\} \\ &\times \sum_{k=0}^{\infty} (\omega_2 2\kappa\sqrt{1 - l^2})^{2k} \frac{1}{4^k} \frac{1}{(k!)^2}, \\ &= l_{(0,1)}(l) l_{(0,\infty)}(\kappa) \frac{1}{(\lambda_1\lambda_2)^{\frac{n}{2}}(n-2)!} l^{n-2} \exp\{-\omega_1 2\kappa\} \\ &\times \sum_{k=0}^{\infty} \frac{1}{(k!)^2} (\omega_2 \sqrt{1 - l^2})^{2k} \kappa^{2k+n-1}. \end{aligned}$$

Es gilt

$$\begin{aligned} &\frac{1}{(\lambda_1\lambda_2)^{\frac{n}{2}}} \omega_2^{2k} (2\omega_1)^{-2k-n+1} \cdot (2\omega_1)^{-1} \\ &= \left(\frac{\sqrt{\lambda_1\lambda_2}}{\frac{1}{2}(\lambda_1 + \lambda_2)} \right)^n \left(\frac{\lambda_1 - \lambda_2}{2(\lambda_1 + \lambda_2)} \right)^{2k} \\ &= \lambda^n (1 - \lambda^2)^k \frac{1}{2^{2k}} \end{aligned}$$

und

$$\begin{aligned} &\int_0^{\infty} \frac{1}{(\lambda_1\lambda_2)^{\frac{n}{2}}} \omega_2^{2k} \kappa^{2k+n-2} \exp\{-\omega_1 2\kappa\} d\kappa \\ &= \lambda^n (1 - \lambda^2)^k \frac{1}{2^{2k}} \int_0^{\infty} (2\omega_1 \kappa)^{2k+n-1} \exp\{-\omega_1 2\kappa\} d(2\omega_1 \kappa) \\ &= \lambda^n (1 - \lambda^2)^k \frac{1}{2^{2k}} (2k + n - 1)!. \end{aligned}$$

Das ergibt

$$f_L(l) = \int_0^\infty f_{L,\kappa}(l, \kappa) d\kappa = l_{(0,1)}(l) \frac{1}{(n-2)!} l^{n-2} \\ \times \sum_{k=0}^\infty \frac{(2k+n-1)!}{(k!)^2 2^{2k}} \lambda^n (1-\lambda^2)^k (1-l^2)^k.$$

Damit ist die behauptete Darstellung von f_L nachgewiesen. Jetzt seien die Eigenwerte gleich. Dann gilt

$$f_L(l) = l_{(0,1)}(l) \frac{1}{(n-2)!} l^{n-2} (n-1)! = l_{(0,1)}(l) (n-1) l^{n-2}.$$

■

Weil wir die Verteilung der Teststatistik sowohl unter der Nullhypothese als auch unter der Alternative kennen, können wir einen Test konstruieren, der exakt das Niveau einhält und gleichzeitig die Gütefunktion exakt berechnen. Wir bezeichnen durch

$$c_n(\gamma) = \gamma^{\frac{1}{n-1}}$$

das Gammaquantil der Dichte f_L in (3.17). Dann ist

$$\varphi_n = l_{(0, \alpha^{\frac{1}{n-1}})}(L)$$

ein α -Test für das Testproblem

$$H_0 : N_{0,\Sigma}^{\otimes n}, \Sigma \in \mathfrak{S}_0 \quad \text{gegen} \quad H_A : N_{0,\Sigma}^{\otimes n}, \Sigma \in \mathfrak{S} \setminus \mathfrak{S}_0.$$

Der Test φ_n wird in der Literatur als Mauchly-Test bezeichnet. Die Wahl des kritischen Bereichs als Bereich, wo L kleine Werte annimmt, ist plausibel, weil $0 \leq L \leq 1$ gilt und $L = 1$ für $L_1 = L_2$ erfüllt ist. Natürlich ergibt sich dieser kritische Bereich auch aus der Konstruktion des Likelihoodquotiententests, siehe Lemma 3.8.

Der hier betrachtete Fall der Dimension zwei ist unter der Nullhypothese nur ein Spezialfall eines allgemeineren Resultats aus der Literatur, siehe Anderson (1984), Kapitel 10.7.3. Liegt eine d -dimensionale Normalverteilung vor und bezeichnen $L_1 \geq L_2 \geq \dots \geq L_d$ die d Eigenwerte der empirischen Kovarianzmatrix, so betrachtet man die Teststatistik

$$L = \frac{\prod_{i=1}^d L_i^{1/d}}{\frac{1}{d} \sum_{i=1}^d L_i}.$$

Unter der Nullhypothese, dass alle Eigenwerte gleich sind, hat L wieder die obige Dichte f_L . Im Gegensatz zum zweidimensionalen Fall gibt es keine expliziten Resultate über die Verteilung der Teststatistik unter der Alternative.

Beispiel 3.13 In diesem Beispiel wollen wir Testergebnisse mit Hilfe einer Simulation darstellen. Dazu erzeugen wir $n = 100$ Zufallszahlen mit der Kovarianzmatrix Σ , die die Eigenwerte λ_1 und λ_2 hat. Die Tabelle 3.1 zeigt die Testergebnisse des Likelihoodquotiententests.

Die Abbildung 3.1 stellt die simulierten Testdaten für unterschiedliche λ_1 und λ_2 dar. Erkennbar ist ein Zunehmen der Korrelation bei wachsendem Abstand der Eigenwerte.

3 Testen auf Hauptrichtungen des Wachstums

Testergebnisse für simulierte Daten					
				Kritische Werte	
			$\alpha = 0.1$	$\alpha = 0.05$	$\alpha = 0.01$
		Teststatistik	$c_n(\alpha) = 0.9112$	$c_n(\alpha) = 0.9413$	$c_n(\alpha) = 0.9545$
Daten 1	$\lambda_1 = 1, \lambda_2 = 1$	0.9919	Annehmen	Annehmen	Annehmen
Daten 2	$\lambda_1 = 1.9, \lambda_2 = 0.1$	0.2735	Ablehnen	Ablehnen	Ablehnen
Daten 3	$\lambda_1 = 3, \lambda_2 = 0.01$	0.0165	Ablehnen	Ablehnen	Ablehnen

Tabelle 3.1: Testergebnisse des Likelihoodquotiententests.

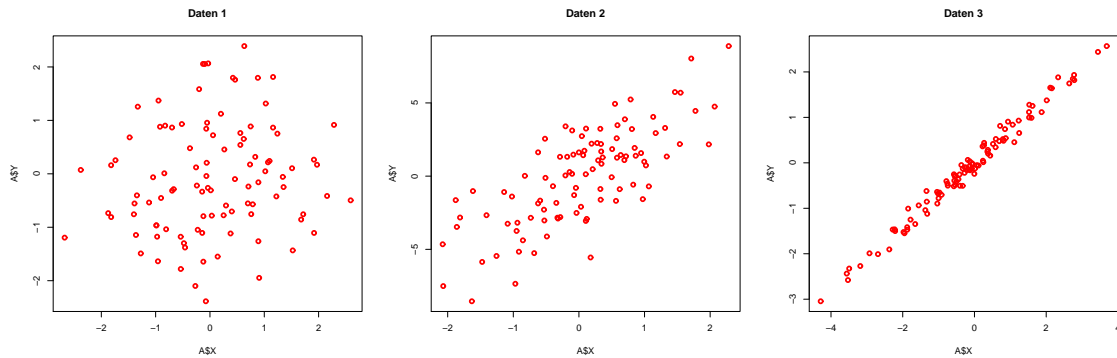


Abbildung 3.1: Simulierte Daten

Wir wollen jetzt die Güte des Mauchly-Tests für die Dimension zwei berechnen. In der nächsten Folgerung geben wir eine explizite Darstellung der Gütefunktion

$$g_M(\alpha, \lambda) = \int_0^{\alpha^{1/(n-1)}} f_L(l, \alpha) dl,$$

die sich unmittelbar aus Satz 3.11 ergibt.

Folgerung 3.14 Die Gütefunktion $g_M(\alpha, \lambda)$ des Mauchly-Tests ist gegeben durch

$$g_M(\alpha, \lambda) = \frac{\lambda^n}{(n-2)!} \sum_{k=0}^{\infty} \frac{(2k+n-1)!}{(k!)^2 2^{2k}} (1-\lambda^2)^k \int_0^{\alpha^{1/(n-1)}} (1-l^2)^k l^{n-2} dl.$$

Die Abbildung 3.2 gibt für $n = 20, 50$ bzw. 100 und für $\alpha = 0.05$ den Verlauf der Gütefunktion an.

Diese Bilder zeigen einen monotonen Verlauf der Gütefunktion. Ein streng mathematischer Beweis für diese Eigenschaft konnte bisher noch nicht erbracht werden. Das Problem ist, dass λ^n monoton wachsend und $(1-\lambda^2)^k$ monoton fallend in λ sind.

Gibt man sich jetzt eine Mindestdifferenz Δ von der Nullhypothese $H_0 : \lambda = 1$ vor, so kann man nach dem notwendigen Stichprobenumfang $n(\alpha, \beta, \Delta)$ fragen, für den bei vorgegebenem $\beta > 0$ die Fehlerwahrscheinlichkeit zweiter Art kleiner oder gleich β ist, wenn λ zur eingeschränkten Alternative $0 < \lambda \leq 1 - \Delta$ gehört. Geht man von einer Monotonie der Gütefunktion aus, dann ergibt sich $N(\alpha, \beta, \Delta, n)$ als minimales n , das der Bedingung

$$\frac{(1-\Delta)^n}{(n-2)!} \sum_{k=0}^{\infty} \frac{(2k+n-1)!}{(k!)^2 2^{2k}} (1-(1-\Delta)^2)^k \int_0^{\alpha^{1/(n-1)}} (1-l^2)^k l^{n-2} dl \geq 1 - \beta$$

genügt.

3 Testen auf Hauptrichtungen des Wachstums

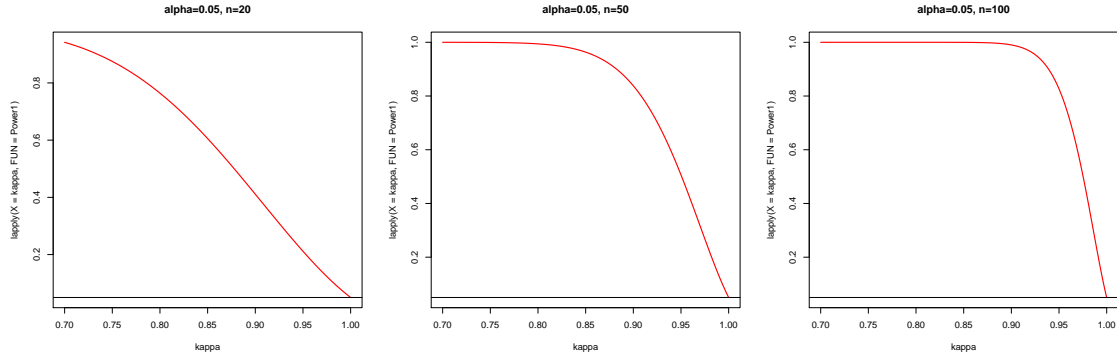


Abbildung 3.2: Verlauf der Gütefunktion für $\alpha = 0.05$

3.2 Asymptotische Tests auf Vorliegen einer Hauptrichtung im Einclustermodell

In diesem Kapitel wollen wir mit Hilfe asymptotischer Tests ohne die Voraussetzung der Normalverteilung eine mögliche Hauptrichtung nachweisen. Da wir auch die Güte der Tests unter lokalen Alternativen untersuchen wollen, nehmen wir an, dass ein Dreiecksschema von Daten vorliegt. Genauer seien für jedes n die $(X_{n,i}, Y_{n,i})$, $i = 1, \dots, n$ unabhängige und identisch verteilte zufällige Vektoren mit Erwartungswert Null und endlichen vierten Momenten. Seien

$$\begin{pmatrix} S_{n,1,1} & S_{n,1,2} \\ S_{n,2,1} & S_{n,2,2} \end{pmatrix} = \begin{pmatrix} \frac{1}{n} \sum_{i=1}^n X_{n,i}^2 & \frac{1}{n} \sum_{i=1}^n X_{n,i} Y_{n,i} \\ \frac{1}{n} \sum_{i=1}^n X_{n,i} Y_{n,i} & \frac{1}{n} \sum_{i=1}^n Y_{n,i}^2 \end{pmatrix},$$

$$L_{n,1(2)} = \frac{1}{2}(S_{n,1,1} + S_{n,2,2}) + (-) \sqrt{\frac{1}{4}(S_{n,1,1} - S_{n,2,2})^2 + S_{n,1,2}^2}$$

die empirische Kovarianzmatrix und ihre Eigenwerte. Es gilt

$$\begin{aligned} L_{n,1} + L_{n,2} &= S_{n,1,1} + S_{n,2,2}, \\ L_{n,1} L_{n,2} &= S_{n,1,1} S_{n,2,2} - S_{n,1,2}^2. \end{aligned}$$

Wir wenden jetzt die Likelihoodquotientenstatistik (3.13) aus dem Normalverteilungsmodell für die allgemeine Situation an. Es gilt

$$n \left(1 - \left(\frac{\sqrt{L_{n,1} L_{n,2}}}{\frac{1}{2}(L_{n,1} + L_{n,2})} \right)^2 \right) = \frac{1}{(S_{n,1,1} + S_{n,2,2})^2} ((\sqrt{n}(S_{n,1,1} - S_{n,2,2}))^2 + (2\sqrt{n}S_{n,1,2})^2). \quad (3.19)$$

Das asymptotische Verhalten dieser Folge von Teststatistiken wurde für beliebige Dimension bereits in Waternaux (1977) für den Fall elliptisch konturierter Verteilungen untersucht. Dort wird auch die asymptotische Verteilung unter lokalen Alternativen abgeleitet. Für den allgemeinen Fall, wo nur die Endlichkeit gewisser Momente gefordert wird, wurden Grenzwertsätzen unter der Nullhypothese in Gupta und Xu (2006) und in der dort zitierten Literatur abgeleitet. Als Grenzverteilungen treten Verteilungen von Linearkombinationen χ^2 -verteilter Zufallsvariablen auf. Dieser Effekt des Abweichens der Grenzverteilung von der χ^2 -Verteilung wird in der Literatur über multivariate Statistik als Nichtrobustheit bezeichnet.

3 Testen auf Hauptrichtungen des Wachstums

Wir wollen jetzt für die Dimension zwei Ergebnisse ableiten, die die von Waternaux (1977) und Gupta und Xu (2006) erzielten Resultate als Spezialfälle enthalten. Weiterhin modifizieren wir die Teststatistik so, dass die neue Teststatistik in dem oben geschilderten Sinne robust ist. In einem ersten Schritt formulieren wir einen geeigneten zentralen Grenzwertsatz. Dazu sind Momentenforderungen notwendig.

Für jedes n sind $(X_{n,i}, Y_{n,i})$, $i = 1, \dots, n$ unabhängig und identisch verteilt. (3.20)

$$\lim_{c \rightarrow \infty} \sup_n E|X_{n,1}|^4 I_{[c,\infty)}(|X_{n,1}|) = 0 \quad \text{und} \quad \lim_{c \rightarrow \infty} \sup_n E|Y_{n,1}|^4 I_{[c,\infty)}(|Y_{n,1}|) = 0, \quad (3.21)$$

$$EX_{n,i} = EY_{n,i} = 0, \quad (3.22)$$

$$\Sigma_n = \begin{pmatrix} \sigma_n^2 & \rho_n \sigma_n \tau_n \\ \rho_n \sigma_n \tau_n & \tau_n^2 \end{pmatrix} = \begin{pmatrix} EX_{n,i}^2 & EX_{n,i} Y_{n,i} \\ EX_{n,i} Y_{n,i} & EY_{n,i}^2 \end{pmatrix}, \quad (3.23)$$

$$\lim_{n \rightarrow \infty} \sqrt{n} \left(\begin{pmatrix} \sigma_n^2 & \rho_n \sigma_n \tau_n \\ \rho_n \sigma_n \tau_n & \tau_n^2 \end{pmatrix} - \begin{pmatrix} \sigma^2 & 0 \\ 0 & \sigma^2 \end{pmatrix} \right) = \begin{pmatrix} \beta_{1,1} & \beta_{1,2} \\ \beta_{2,1} & \beta_{2,2} \end{pmatrix}. \quad (3.24)$$

Sei

$$\Gamma_n = \text{Kovarianzmatrix von} \begin{pmatrix} X_{n,i}^2 - Y_{n,i}^2 \\ 2X_{n,i} Y_{n,i} \end{pmatrix} \text{ mit} \\ \lim_{n \rightarrow \infty} \Gamma_n = \begin{pmatrix} \gamma_{1,1} & \gamma_{1,2} \\ \gamma_{2,1} & \gamma_{2,2} \end{pmatrix} = \Gamma, \quad \det(\Gamma) \neq 0. \quad (3.25)$$

Lemma 3.15 *Es seien die Voraussetzungen (3.20) bis (3.25) erfüllt. Dann konvergieren die Verteilungen der zufälligen Vektoren*

$$V_n = \sqrt{n} (S_{n,1,1} - S_{n,2,2}, 2S_{n,1,2})^T \quad (3.26)$$

schwach gegen eine Normalverteilung mit dem Erwartungswertvektor

$$\mu = (\beta_{1,1} - \beta_{2,2}, 2\beta_{1,2})^T$$

und der Kovarianzmatrix Γ in (3.25).

Beweis. Seien $v = (a_1, a_2)^T \in \mathbb{R}^2$ ein fester Vektor und

$$V_{n,j} = (X_{n,j}^2 - Y_{n,j}^2 - (\sigma_n^2 - \tau_n^2), X_{n,j} Y_{n,j} - \rho_n \tau_n \sigma_n)^T.$$

Weiter seien

$$\begin{aligned} Z_{n,j} &= a_1(X_{n,j}^2 - Y_{n,j}^2 - (\sigma_n^2 - \tau_n^2)) + a_2(2X_{n,j} Y_{n,j} - \rho_n \tau_n \sigma_n), \\ b_n^2 &= V(Z_{n,j}). \end{aligned}$$

Aus den Voraussetzungen (3.21) ergibt sich

$$\lim_{c \rightarrow \infty} \sup_n E|Z_{n,1}|^2 I_{[c,\infty)}(|Z_{n,1}|) = 0.$$

Die Verwendung von (3.25) ergibt

$$b_n^2 \rightarrow b^2 = v^T \Gamma v.$$

3 Testen auf Hauptrichtungen des Wachstums

Mit Hilfe des Satzes von Cramer-Wold, siehe z. B. Kallenberg (1997) und Lemma 5.11 folgt

$$\mathcal{L}(\sqrt{n}(S_{n,1,1} - S_{n,2,2} - (\sigma_n^2 - \tau_n^2), 2S_{n,1,2} - 2\rho_n\tau_n\sigma_n)^T) \Rightarrow N_{0,\Gamma}.$$

Die Behauptung des Lemmas ergibt sich hieraus und aus der Voraussetzung (3.24). ■

Weil der Faktor $(S_{n,1,1} + S_{n,2,2})^{-2}$ in (3.19) gegen eine Konstante strebt, können wir auch mit der modifizierte Teststatistik

$$T_n = (\sqrt{n}(S_{n,1,1} - S_{n,2,2}))^2 + (2\sqrt{n}S_{n,1,2})^2 = V_n^T V_n \quad (3.27)$$

arbeiten. T_n verhält sich nach Lemma 3.15 in Verteilung asymptotisch wie das Längenquadrat eines normalverteilten Vektors, der aber im Allgemeinen keine Diagonalmatrix als Kovarianzmatrix hat, und man kann deshalb keine χ^2 -Verteilung als Grenzverteilung erwarten. Der nächste Satz zeigt, dass dann die Grenzverteilung von T_n eine Linearkombination von χ^2 -verteilten Zufallsvariablen ist. Zur Formulierung des Ergebnisses benötigen wir noch zusätzliche Bezeichnungen. Es seien

$$\kappa_{1(2)} = \frac{1}{2}(\gamma_{2,2}^2 + \gamma_{1,1}^2) + (-)\frac{1}{2}\sqrt{(\gamma_{2,2}^2 - \gamma_{1,1}^2)^2 + 4\gamma_{1,2}^2} \quad (3.28)$$

die Eigenwerte der Matrix Γ und $e_{1,\Gamma}, e_{2,\Gamma}$ die zugehörigen Eigenvektoren. Um eine Grenzverteilung mit einfacher Struktur zu erhalten normieren wir V_n mit der Kovarianzmatrix Γ und betrachten die Teststatistik $Q_n = V_n^T \Gamma^{-1} V_n$. In den meisten Anwendungen sind die in die Matrix Γ eingehenden vierten Momente nicht bekannt und müssen aus den Daten geschätzt werden. Unter der Voraussetzung (3.24) gilt

$$\lim_{n \rightarrow \infty} E(X_{n,i}^2 - Y_{n,i}^2) = 0 \quad \text{und} \quad \lim_{n \rightarrow \infty} E(2X_{n,i}Y_{n,i}) = 0.$$

Deshalb können wir folgende Schätzer verwenden

$$\begin{aligned} \hat{\gamma}_{n,1,1} &= \frac{1}{n} \sum_{i=1}^n (X_{n,i}^2 - Y_{n,i}^2)^2 \quad \text{und} \quad \hat{\gamma}_{n,2,2} = \frac{4}{n} \sum_{i=1}^n (X_{n,i}Y_{n,i})^2, \\ \hat{\gamma}_{n,1,2} &= \hat{\gamma}_{n,2,1} = \frac{2}{n} \sum_{i=1}^n (X_{n,1}^3 Y_{n,1} - X_{n,1} Y_{n,1}^3), \\ \hat{\Gamma}_n &= \begin{pmatrix} \hat{\gamma}_{n,1,1} & \hat{\gamma}_{n,1,2} \\ \hat{\gamma}_{n,2,1} & \hat{\gamma}_{n,2,2} \end{pmatrix}. \end{aligned}$$

Wir schätzen die Eigenwerte durch

$$\hat{\kappa}_{n,1(2)} = \frac{1}{2}(\hat{\gamma}_{n,2,2}^2 + \hat{\gamma}_{n,1,1}^2) + (-)\frac{1}{2}\sqrt{(\hat{\gamma}_{n,2,2}^2 - \hat{\gamma}_{n,1,1}^2)^2 + 4\hat{\gamma}_{n,1,2}^2} \quad (3.29)$$

und betrachten die modifizierte Teststatistik

$$\hat{Q}_n = V_n^T \hat{\Gamma}_n^{-1} V_n. \quad (3.30)$$

Satz 3.16 *Unter den Voraussetzungen (3.20) bis (3.25) gilt*

$$\begin{aligned} \mathcal{L}(T_n) &\implies \mathcal{L}(\kappa_1(Z_1 + \kappa_1^{-1/2} e_{1,\Gamma}^T \mu)^2 + \kappa_2(Z_2 + \kappa_2^{-1/2} e_{2,\Gamma}^T \mu)^2), \\ \mathcal{L}(Q_n) &\implies \text{CH}_{2,\delta^2} \quad \text{und} \quad \mathcal{L}(\hat{Q}_n) \implies \text{CH}_{2,\delta^2}. \end{aligned}$$

Hierbei sind Z_1, Z_2 unabhängige und standardnormalverteilte Zufallsvariablen. CH_{2,δ^2} ist die Chiquadratverteilung mit zwei Freiheitsgraden und dem Nichtzentralitätsparameter $\delta^2 = \mu^T \Gamma^{-1} \mu$ mit $\mu = (\beta_{1,1} - \beta_{2,2}, 2\beta_{1,2})^T$.

Beweis. Der Beweis der ersten beiden Aussagen folgt aus Lemma 3.15 und Lemma 5.13. Die zweite Behauptung folgt aus Lemma 3.15 und Lemma 5.14. Zum Beweis der dritten Aussage bemerken wir, dass aus dem Gesetz der großen Zahlen für ein Serienschema, siehe Lemma 5.12 und der Voraussetzung (3.21) folgt $\hat{\Gamma}_n \xrightarrow{P} \Gamma$. Weil die Vektoren $(\sqrt{n}(S_{n,1,1} - S_{n,2,2}), 2\sqrt{n}S_{n,1,2})$ nach Lemma 3.15 stochastisch beschränkt sind, können wir nach dem Lemma von Slutsky asymptotisch Γ^{-1} durch $(\hat{\Gamma}_n)^{-1}$ ersetzen. Damit folgt die dritte Behauptung aus der zweiten Aussage. ■

Wir spezialisieren jetzt die betrachteten Modelle und Hypothesen. Um festzustellen, ob sich durch Versuchsänderungen neue Hauptrichtungen herausgeprägt haben, nehmen wir an, dass die beobachteten Daten die Verteilung $\mathcal{L}(A(X, Y)^T)$ haben und wir wollen überprüfen, ob A weder aus einer reinen Streckung noch aus einer Drehung besteht. Das folgende einfache Lemma klärt den Zusammenhang zu den Eigenwerten von AA^T .

Lemma 3.17 *Sei I die Einheitsmatrix. Für eine beliebige reguläre quadratische Matrix A sind folgende Aussagen gleichwertig*

- 1) $A = \gamma O$ mit einer orthogonalen Matrix O und einer reellen Zahl $\gamma \neq 0$,
- 2) $AA^T = \sigma^2 I$,
- 3) AA^T hat zwei gleiche Eigenwerte.

Beweis. Die Aussage 1) \rightarrow 2) \rightarrow 3) ist klar. Gilt 3), dann gibt es für die positiv semidefinite Matrix AA^T eine orthogonale Matrix O mit der Eigenschaft

$$O^T AA^T O = \lambda I.$$

Dann ist die Matrix $\tilde{O} = \lambda^{-1/2} O^T A$ orthogonal und es gilt

$$A = \lambda^{1/2} O \tilde{O}.$$

Die Aussage 1) folgt mit $\gamma = \lambda^{1/2}$ aus der Tatsache, dass $O\tilde{O}$ wieder eine orthogonale Matrix ist. ■

Es sei Q_A die zu

$$f_A(x, y) = (\det(A))^{-1} f(A^{-1}(x, y)^T)$$

gehörige Verteilung, wobei wir die Existenz vierter Momente voraussetzen, d. h.

$$\int \left(\int (x^4 + y^4) f(x, y) dx \right) dy < \infty. \quad (3.31)$$

Weiterhin gelte

$$\begin{aligned} \int \left(\int x f(x, y) dx \right) dy &= \int \left(\int y f(x, y) dx \right) dy = \int \left(\int xy f(x, y) dx \right) dy = 0, \\ \int \left(\int x^2 f(x, y) dx \right) dy &= \int \left(\int y^2 f(x, y) dx \right) dy = 1. \end{aligned} \quad (3.32)$$

Dann ist AA^T die zu Q_A gehörige Kovarianzmatrix. Weiterhin sei

$$\Gamma = \begin{pmatrix} \int \left(\int (x^2 - y^2)^2 f(x, y) dx \right) dy & \int \left(\int 2xy(x^2 - y^2) f(x, y) dx \right) dy \\ \int \left(\int 2xy(x^2 - y^2) f(x, y) dx \right) dy & \int \left(\int (2xy)^2 f(x, y) dx \right) dy \end{pmatrix}. \quad (3.33)$$

3 Testen auf Hauptrichtungen des Wachstums

Bei einer Messung vom Umfang n liegt also folgendes Modell und folgendes Testproblem vor.

$$(\mathbb{R}_{2n}, \mathfrak{B}_{2n}, Q_A^{\otimes n}), \quad A \in \mathfrak{M}, \quad f \text{ erfüllt (3.31) und (3.32)}, \quad (3.34)$$

$$\text{Nullhypothese: } H_0 : \Sigma = AA^T \in \mathfrak{S}_0,$$

$$\text{Alternativhypothese: } H_A : \Sigma = AA^T \in \mathfrak{S}_A = \mathfrak{S} \setminus \mathfrak{S}_0.$$

Zur Untersuchung der lokalen Güte betrachten wir die lokale Alternative

$$\text{Lokale Alternative: } H_{A,loc} : A_n = \sigma I + \frac{1}{\sqrt{n}}A \quad \text{und somit}$$

$$\Sigma_n = A_n A_n^T =: \sigma^2 I + \frac{1}{\sqrt{n}}B + o(n^{-1/2}) \quad \text{mit} \quad B = \begin{pmatrix} \beta_{1,1} & \beta_{1,2} \\ \beta_{2,1} & \beta_{2,2} \end{pmatrix}, \quad \beta_{1,2} = \beta_{2,1}. \quad (3.35)$$

Unter diesen lokalen Alternativen sind die Voraussetzungen (3.20),..., (3.25) automatisch erfüllt.

Lemma 3.18 Seien die $(X_i, Y_i)^T$ u.i.v. mit der Dichte f , die den Bedingungen (3.31) und (3.32) genügt. Mit A_n aus (3.35) erfüllen die $(X_{n,i}, Y_{n,i})^T = A_n(X_i, Y_i)^T$ die Voraussetzungen (3.20) bis (3.25) und es gilt für $0 \leq k + l \leq 4$

$$\frac{1}{n} \sum_{i=1}^n X_{n,i}^k Y_{n,i}^l \rightarrow^P \sigma^{k+l} E(X^k Y^l).$$

Beweis. Die Behauptung ergibt sich aus dem Gesetz der großen Zahlen und der Tatsache, dass $X_{n,i}^k Y_{n,i}^l$ eine Linearkombination von gewissen $X_{n,i}^r Y_{n,i}^m$ ist und die Gewichtungsfaktoren konvergieren. ■

Zur Konstruktion von Tests bezeichnen wir durch $c(\kappa_1, \kappa_2, \gamma)$ das γ -Quantil der Verteilungsfunktion

$$K_{\kappa_1, \kappa_2}(t) = P(\kappa_1 Z_1^2 + \kappa_2 Z_2^2 \leq t),$$

wobei Z_1, Z_2 unabhängig und standardnormal sind. Die Dichte dieser Verteilung wurde bereits im Satz 2.22 untersucht. Wir bemerken, dass $c(1, 1, \gamma)$ das γ -Quantil einer χ^2 -Verteilung mit zwei Freiheitsgraden ist, die die Dichte

$$h_2(t) = I_{[0, \infty)}(t) \frac{1}{2} \exp\left\{-\frac{1}{2}t\right\}$$

besitzt. Damit folgt

$$c(1, 1, \gamma) = -2 \ln(1 - \gamma).$$

Die Quantile $c(\kappa_1, \kappa_2, \gamma)$ für beliebige κ_1, κ_2 kann man numerisch mit Hilfe der Reihendarstellung im Satz 2.22 ermitteln. Wir nehmen jetzt an, dass die Statistiken T_n aus (3.27) und \hat{Q}_n aus (3.30) auf dem Wahrscheinlichkeitsraum in (3.34) definiert sind. Wir führen jetzt zwei Testfolgen ein

$$\varphi_n = I_{[c(\hat{\kappa}_{n,1}, \hat{\kappa}_{n,2}, 1-\alpha), \infty)}(T_n) \quad \text{und} \quad \psi_n = I_{[-2 \ln(\alpha), \infty)}(\hat{Q}_n).$$

Satz 3.19 Es seien die Momentenbedingung (3.31) und die Voraussetzungen in (3.32) erfüllt und Γ sei durch (3.33) definiert.

1. Unter der Nullhypothese H_0 in (3.34) sind die Folgen $\{\varphi_n\}$ und $\{\psi_n\}$ asymptotische α -Tests, d. h.

$$\lim_{n \rightarrow \infty} \int \varphi_n dQ_A^{\otimes n} = \lim_{n \rightarrow \infty} \int \psi_n dQ_A^{\otimes n} = \alpha, \quad AA^T \in \mathfrak{S}_0.$$

3 Testen auf Hauptrichtungen des Wachstums

2. Unter der lokalen Alternative $H_{A,loc}$ in (3.35) haben die Folgen von Tests $\{\varphi_n\}$ die asymptotischen Güten

$$\begin{aligned}\lim_{n \rightarrow \infty} \int \varphi_n dQ_{A_n}^{\otimes n} &= g_1(\Gamma, \alpha, \mu), \\ \lim_{n \rightarrow \infty} \int \psi_n dQ_{A_n}^{\otimes n} &= g_2(\alpha, \mu^T \Gamma^{-1} \mu)\end{aligned}$$

mit

$$\begin{aligned}g_1(\Gamma, \alpha, \mu) &= P(\kappa_1(Z_1 + \kappa_1^{-1/2} e_{1,\Gamma}^T \mu)^2 + \kappa_2(Z_2 + \kappa_2^{-1/2} e_{2,\Gamma}^T \mu)^2) > c(1 - \alpha, \kappa_1, \kappa_2), \\ g_2(\alpha, \mu^T \Gamma^{-1} \mu) &= CH_{2,\delta^2}((-2 \ln(\alpha), \infty)),\end{aligned}$$

wobei $Z_1, Z_2, \mu = (\beta_{1,1} - \beta_{2,2}, 2\beta_{1,2})^T$, CH_{2,δ^2} und δ^2 die gleiche Bedeutung wie im Satz 3.16 haben und $c(\gamma, \kappa_1, \kappa_2)$ das γ -Quantil in (2.52) ist.

Beweis. Die Aussage für ψ_n folgen aus Satz 3.16. Für $\varepsilon > 0$ setzen wir

$$B_{n,\varepsilon} = \{|c(\hat{\kappa}_{n,1}, \hat{\kappa}_{n,2}, 1 - \alpha) - c(\kappa_1, \kappa_2, 1 - \alpha)| > \varepsilon\}.$$

Wegen Lemma 3.18 gilt $Q_A^{\otimes n}(\bar{B}_{n,\varepsilon}) \rightarrow 0$ unter der Nullhypothese. Weiterhin gilt

$$Q_A^{\otimes n}(T_n > c(\hat{\kappa}_{n,1}, \hat{\kappa}_{n,2}, 1 - \alpha)) \leq Q_A^{\otimes n}(T_n > c(\kappa_1, \kappa_2, 1 - \alpha) - \varepsilon) + Q_A^{\otimes n}(\bar{B}_{n,\varepsilon}).$$

Die Stetigkeit von $c(1 - \alpha, \kappa_1, \kappa_2)$ als Funktion von (κ_1, κ_2) , siehe Lemma 2.23, und

$$Q_A^{\otimes n}(\bar{B}_{n,\varepsilon}) \rightarrow 0 \text{ für } n \rightarrow \infty$$

ergeben

$$\begin{aligned}\limsup_{n \rightarrow \infty} Q_A^{\otimes n}(T_n > c(\hat{\kappa}_{n,1}, \hat{\kappa}_{n,2}, 1 - \alpha)) \\ \leq Q_A^{\otimes n}(\kappa_1(Z_1 + \kappa_1^{-1/2} e_{1,\Gamma}^T \mu)^2 + \kappa_2(Z_2 + \kappa_2^{-1/2} e_{2,\Gamma}^T \mu)^2 \\ \geq c(\kappa_1, \kappa_2, 1 - \alpha) - \varepsilon).\end{aligned}$$

Lässt man ε gegen Null streben und schätzt man analog den unteren Limes ab, so folgen die Aussagen für φ_n aus Satz 3.16. ■

Folgerung 3.20 Sind die Eigenwerte κ_1, κ_2 gleich, so sind die asymptotischen Gütefunktionen gleich, die beiden Testfolgen sind also asymptotisch gleichwertig.

Beweis. Zum Beweis der Folgerung setzen wir $\kappa = \kappa_1 = \kappa_2$ und bemerken

$$c(\kappa, \kappa, 1 - \alpha) = \kappa c(1, 1, 1 - \alpha) = -2\kappa \ln \alpha.$$

Deshalb gilt

$$\begin{aligned}\kappa_1(Z_1 + \kappa_1^{-1/2} e_{1,\Gamma}^T \mu)^2 + \kappa_2(Z_2 + \kappa_2^{-1/2} e_{2,\Gamma}^T \mu)^2 &\geq c(\kappa_1, \kappa_2, 1 - \alpha) \\ \iff (Z_1 + \kappa^{-1/2} e_{1,\Gamma}^T \mu)^2 + (Z_2 + \kappa^{-1/2} e_{2,\Gamma}^T \mu)^2 &\geq -2 \ln \alpha.\end{aligned}$$

3 Testen auf Hauptrichtungen des Wachstums

Die Verteilung der linken Seite ist eine χ^2 -Verteilung mit zwei Freiheitsgraden und dem Nichtzentralitätsparameter

$$\delta^2 = (\kappa^{-1/2} e_{1,\Gamma}^T \mu)^2 + (\kappa^{-1/2} e_{2,\Gamma}^T \mu)^2 = \frac{1}{\kappa} \mu^T \mu = \mu^T \Gamma^{-1} \mu.$$

■

Wir betrachten jetzt die homogenen Dichten

$$h_{g,p}(x, y) = C(p) g(|x|^p + |y|^p)$$

in (2.19), überprüfen die Bedingung (3.32) und berechnen die Matrix Γ . Um die Existenz der vierten Momente zu sichern, setzen wir nach Lemma 2.12 voraus

$$\int_0^\infty s^{\frac{6-p}{p}} g(s) ds < \infty. \quad (3.36)$$

Jetzt betrachten wir zunächst den Spezialfall einer verallgemeinerten Normalverteilung, d. h.

$$h_{g_{N,p}}(x_0, y_0) = \frac{1}{\Gamma\left(\frac{2}{p}\right)} \exp\{-|x_0|^p - |y_0|^p\}.$$

Dann gelten

$$\begin{aligned} EX_0 &= EY_0 = EX_0 Y_0 = 0, \\ EX_0^2 &= EY_0^2. \end{aligned}$$

Wir berechnen $\gamma_1(p) = V(X_0^2 - Y_0^2)$ und $\gamma_2(p) = V(2X_0 Y_0)$.

$$\begin{aligned} \gamma_1(p) &= E(X_0^2 - Y_0^2)^2 \\ &= E(X_0^4 - 2X_0^2 Y_0^2 + Y_0^4) \\ &= EX_0^4 - 2EX_0^2 EY_0^2 + EY_0^4 \\ &= \frac{4\Gamma(5/p) - \Gamma(3/p)}{p\Gamma(2/p)}, \\ \gamma_2(p) &= E(2X_0 Y_0)^2 = 4(EX_0^2) = \frac{8\Gamma(3/p)}{p\Gamma(2/p)}. \end{aligned}$$

Wegen

$$\text{cov}(X_0^2 - Y_0^2, 2X_0 Y_0) = E(2X_0 Y_0 X_0^2 - 2X_0 Y_0 Y_0^2) = 0$$

erhalten wir

$$\begin{aligned} \Gamma_0 &= \begin{pmatrix} V(X_0^2 - Y_0^2) & \text{cov}(X_0^2 - Y_0^2, 2X_0 Y_0) \\ \text{cov}(X_0^2 - Y_0^2, 2X_0 Y_0) & V(2X_0 Y_0) \end{pmatrix} \\ &= \begin{pmatrix} \gamma_1(p) & 0 \\ 0 & \gamma_2(p) \end{pmatrix}. \end{aligned}$$

Ist jetzt (X, Y) ein Vektor mit der Dichte $h_{g,p}$ und ist (3.36) erfüllt, so folgt aus Folgerung 2.13, dass (X, Y) auch der Bedingung (3.32) genügt und

$$\begin{aligned} \Gamma_{g,p} &= \begin{pmatrix} V(X^2 - Y^2) & \text{cov}(X^2 - Y^2, 2XY) \\ \text{cov}(X^2 - Y^2, 2XY) & V(2XY) \end{pmatrix} \\ &= \frac{\int_0^\infty s^{\frac{6}{p}-1} g(s) ds}{\int_0^\infty s^{\frac{6}{p}-1} g_{N,p}(s) ds} \begin{pmatrix} \gamma_1(p) & 0 \\ 0 & \gamma_2(p) \end{pmatrix} \end{aligned}$$

gilt.

Aufgrund der speziellen Struktur der Matrix $\Gamma_{g_{N,p}}$ lassen sich die Schätzungen für diese Matrix Γ_0 vereinfachen. Weil

$$\Gamma_{g_{N,p}} = \frac{1}{\int_0^\infty s^{\frac{6}{p}-1} g_{N,p}(s) ds} \begin{pmatrix} \gamma_1(p) & 0 \\ 0 & \gamma_2(p) \end{pmatrix}$$

bekannt ist, wird nur noch die unbekannte Konstante $\int_0^\infty s^{\frac{6}{p}-1} g(s) ds = EL^4$ benötigt. Diese können wir direkt schätzen. Aus $L = R \cdot h^{1/p}(\Phi)$ erhalten wir

$$EL^4 = ER^4 (|\cos \Phi|^p + |\sin \Phi|^p)^{\frac{4}{p}}.$$

Es seien (X_i, Y_i) , $i = 1, \dots, n$ u.i.v. mit der Dichte $h_{g,p}$. Wir setzen

$$R_i = (X_i^2 + Y_i^2)^{1/2}, \quad \Phi_i = \mathcal{P}_2(X_i, Y_i)$$

und

$$\hat{M}_{n,p} = \frac{1}{n} \sum_{i=1}^n R_i^4 (|\cos \Phi_i|^p + |\sin \Phi_i|^p)^{\frac{4}{p}}.$$

Dann gilt nach dem Gesetz der großen Zahlen $M_n \rightarrow EL^4$ und wir können als vereinfachten Schätzen für $\Gamma_{g,p}$ verwenden

$$\hat{\Gamma}_n = \hat{M}_n \cdot \Gamma_{g_{N,p}}.$$

3.3 Asymptotische Tests auf Vorliegen einer Hauptrichtung im Mehrclustermmodell

In diesem Abschnitt wollen wir einen asymptotischen Test konstruieren, wenn die Voraussetzungen im Modell 2.2 erfüllt sind. Die beobachteten zufälligen Vektoren $(X_1, Y_1), \dots, (X_n, Y_n)$ haben den Erwartungswert Null und Kovarianzmatrizen. Wir nehmen zusätzlich an, dass die Vektoren normalverteilt sind. Genauer möge es gewisse orthogonalen Matrizen O_j derart geben, dass die Vektoren $(U_j, V_j)^T = O_j^T (X_j, Y_j)^T$ unabhängig sind und die gleiche Normalverteilung besitzen. Dann sind auch die

$$R_j^2 = U_j^2 + V_j^2 = X_j^2 + Y_j^2 \tag{3.37}$$

unabhängig und identisch verteilt. Die Verteilung von R_j^2 hatten wir im Satz 2.20 ermittelt, wobei $\mathcal{L}(X_j, Y_j) \in \mathfrak{E}$ vorausgesetzt wurde. Betrachtet man speziell die Familie von Dichten $\kappa_{\lambda_1, \lambda_2}$, so erkennt man, dass die Parameter λ_1, λ_2 durch die ausschließliche Beobachtung von R_1^2, \dots, R_n^2 nicht identifiziert werden können. Durch Vertauschen von λ_1 und λ_2 ergibt sich die gleiche Verteilung. Schränkt man den Parameterbereich durch die Forderung $\lambda_1 \geq \lambda_2$ ein, so liegt die Nullhypothese am Rand des Parameterbereichs, was in den klassischen Sätzen über den Likelihoodquotiententest ausgeschlossen ist. Deshalb werden wir eine andere geeignete Teststatistik konstruieren und deren asymptotisches Verhalten untersuchen. Weil das Testproblem $H_0 : \lambda_1 = \lambda_2$ gegen $H_A : \lambda_1 > \lambda_2$ invariant gegenüber Skalentransformationen ist, verwenden wir die folgende skaleninvariante Teststatistik

$$T_n = \frac{(\frac{1}{n} \sum_{i=1}^n R_i^2)^2}{\frac{1}{n} \sum_{i=1}^n R_i^4}.$$

3 Testen auf Hauptrichtungen des Wachstums

Die Zufallsvariablen U_j, V_j seien unabhängig und standardnormal. Weil die Vektoren (X_j, Y_j) normalverteilt sind mit Erwartungswert Null und der Kovarianzmatrix Σ mit den Eigenwerten λ_1, λ_2 , gelten

$$\mathcal{L}(R_j^2) = \mathcal{L}(X_j^2 + Y_j^2) = \mathcal{L}(\lambda_1 U_j^2 + \lambda_2 V_j^2)$$

und

$$\begin{aligned} \mathcal{L}(T_n) &= \mathcal{L}(S_n) \text{ mit} \\ S_n &= \frac{\left(\frac{1}{n} \sum_{i=1}^n (\lambda_1 U_i^2 + \lambda_2 V_i^2)\right)^2}{\frac{1}{n} \sum_{i=1}^n (\lambda_1 U_i^2 + \lambda_2 V_i^2)^2}. \end{aligned}$$

Aus dem Gesetz der großen Zahlen erhalten wir deshalb

$$T_n \xrightarrow{P} \psi(\lambda_1, \lambda_2) := \frac{(\lambda_1 + \lambda_2)^2}{3\lambda_1^2 + 3\lambda_2^2 + 2\lambda_1\lambda_2} = \frac{(1 + \kappa)^2}{3 + 3\kappa^2 + 2\kappa} =: g(\kappa),$$

wobei $\kappa = \frac{\lambda_2}{\lambda_1}$ gesetzt wurde. Wir untersuchen die Funktion g näher. Es gilt

$$g'(\kappa) = 4(1 + \kappa) \frac{1 - \kappa}{(3 + 3\kappa^2 + 2\kappa)^2}.$$

Dann gilt für alle $\kappa \neq 1$

$$g(\kappa) < g(1) = \frac{1}{2}.$$

Deshalb und wegen $T_n \xrightarrow{P} g(\kappa)$ bietet sich

$$W_n := \sqrt{n} \left(T_n - \frac{1}{2} \right) \quad (3.38)$$

als Teststatistik an. Aus (3.37) folgt, dass die Verteilung von W_n nur vom Quotienten der Eigenwerte der Kovarianzmatrix Σ abhängt. Deshalb können wir ohne Einschränkung der Allgemeinheit zu folgender speziellen Darstellung von T_n übergehen

$$\begin{aligned} T_n &= \frac{\left(\frac{1}{n} \sum_{i=1}^n (U_i^2 + \kappa_n V_i^2)\right)^2}{\frac{1}{n} \sum_{i=1}^n (U_i^2 + \kappa_n V_i^2)^4} \\ &= \frac{\left(\frac{1}{n} \sum_{i=1}^n R_{i,n}^2\right)^2}{\frac{1}{n} \sum_{i=1}^n R_{i,n}^4}, \end{aligned}$$

$$\text{wobei } U_1, \dots, U_n, V_1, \dots, V_n \text{ u.i.v. standardnormal.} \quad (3.39)$$

In dieser Darstellung modellieren die Konstanten $\kappa_n > 0$ mögliche Abweichungen von der Nullhypothese $H_0 : \kappa = 1$. Zur Untersuchung des asymptotischen Verhaltens von T_n stellen wir zunächst Hilfsresultate bereit.

Lemma 3.21 *U, V seien unabhängige und standardnormalverteilte Zufallsvariablen. Dann besitzt $Z = \frac{1}{2}(U^2 + V^2)$ eine Standardexponentialverteilung mit der Dichte $f(x) = I_{[0, \infty)}(x) \exp\{-x\}$. Es gilt*

$$E(U^2 + V^2)^k = 2^k k!. \quad (3.40)$$

Weiterhin gelten für $Z_\kappa = U^2 + \kappa V^2$ und $j = 1, 2, \dots$

$$\begin{aligned} E Z_\kappa &= 1 + \kappa, \\ V(Z_\kappa) &= 2(1 + \kappa^2), \\ E(Z_\kappa)^j &\leq \frac{(1 + \kappa^j)(2j)!}{2j!}. \end{aligned} \quad (3.41)$$

3 Testen auf Hauptrichtungen des Wachstums

Beweis. $U^2 + V^2$ besitzt eine χ^2 -Verteilung mit zwei Freiheitsgraden und damit die Dichte $h_2(x) = I_{[0,\infty)}(x) \frac{1}{2} \exp\{-x/2\}$. Das beweist die erste Behauptung. Hat Z die Exponentialverteilung $I_{[0,\infty)}(x) \exp\{-x\}$, dann gilt

$$EZ^k = k!$$

und somit (3.40). Weiterhin gelten wegen $EU^2 = EV^2 = 1$ und $EU^4 = EV^4 = 3$

$$\begin{aligned} EZ_{\kappa} &= 1 + \kappa, \\ EZ_{\kappa}^2 &= 3 + 2\kappa + 3\kappa^2, \\ V(Z_{\kappa}) &= EZ_{\kappa}^2 - (EZ_{\kappa})^2 \\ &= 2(1 + \kappa^2). \end{aligned}$$

Die behauptete Ungleichung folgt aus $(a + b)^j \leq 2^{j-1}(a^j + b^j)$ und

$$EU^{2j} = EV^{2j} = \frac{(2j)!}{2^j j!}.$$

■

Wir formen W_n um

$$\begin{aligned} \sqrt{n} \left(T_n - \frac{1}{2} \right) &= \sqrt{n} \left(\frac{\left(\frac{1}{n} \sum_{i=1}^n R_{i,n}^2 \right)^2}{\frac{1}{n} \sum_{i=1}^n R_{i,n}^4} - \frac{1}{2} \right) \\ &= \frac{1}{\frac{1}{n} \sum_{i=1}^n R_{i,n}^4} \left[\sqrt{n} \left(\left(\frac{1}{n} \sum_{i=1}^n R_{i,n}^2 \right)^2 - 4 \right) - \frac{1}{2\sqrt{n}} \sum_{i=1}^n (R_{i,n}^4 - 8) \right]. \end{aligned}$$

Wir untersuchen zunächst den ersten Term in der Klammer.

Lemma 3.22 *Unter den Voraussetzungen (3.39), $\kappa_n > 0$ und*

$$\lim_{n \rightarrow \infty} \sqrt{n}(\kappa_n - 1) = h \quad (3.42)$$

gilt

$$\sqrt{n} \left(\left(\frac{1}{n} \sum_{i=1}^n R_{i,n}^2 \right)^2 - 4 \right) = \frac{4}{\sqrt{n}} \sum_{i=1}^n (R_{i,n}^2 - 2) + o_P(1).$$

Beweis. Es gilt

$$\begin{aligned} &\sqrt{n} \left(\left(\frac{1}{n} \sum_{i=1}^n R_{i,n}^2 \right)^2 - 4 \right) \\ &= \left(\frac{1}{\sqrt{n}} \sum_{i=1}^n (R_{i,n}^2 - 2) \right) \left(\left(\frac{1}{n} \sum_{i=1}^n R_{i,n}^2 \right) + 2 \right). \end{aligned}$$

Weiterhin gelten

$$\begin{aligned} ER_{1,n}^2 &= E(U_1^2 + \kappa_n V_1^2) = 1 + \kappa_n \rightarrow 2, \\ ER_{1,n}^4 &= E(U_1^2 + \kappa_n V_1^2)^2 = EU_1^4 + \kappa_n^2 EV_1^4 + 2\kappa_n \rightarrow 8. \end{aligned}$$

3 Testen auf Hauptrichtungen des Wachstums

Aus (3.41) folgt

$$V\left(\frac{1}{n}\sum_{i=1}^n R_{i,n}^2\right) = \frac{1}{n}V(R_{1,n}^2) = \frac{1}{n}(ER_{1,n}^4 - (ER_{1,n}^2)^2) \rightarrow 0.$$

Damit gilt

$$E\left(\frac{1}{n}\sum_{i=1}^n R_{i,n}^2 - 2\right)^2 = V\left(\frac{1}{n}\sum_{i=1}^n R_{i,n}^2\right) + (1 + \kappa_n - 2)^2 \rightarrow 0$$

und es folgt

$$\frac{1}{n}\sum_{i=1}^n R_{i,n}^2 = 2 + o_P(1).$$

Weiterhin erhalten wir aus (3.41)

$$\begin{aligned} E\left(\frac{1}{\sqrt{n}}\sum_{i=1}^n (R_{i,n}^2 - 2)\right)^2 &= V(R_{1,n}^2) + (1 + \kappa_n - 2)^2 \\ &= 2(1 + \kappa_n^2) + (1 + \kappa_n - 2)^2 \rightarrow 4. \end{aligned} \quad (3.43)$$

Damit bleibt die linke Seite beschränkt und somit ist $\frac{1}{\sqrt{n}}\sum_{i=1}^n (R_{i,n}^2 - 2)$ stochastisch beschränkt. Das ergibt

$$\begin{aligned} &\left(\frac{1}{\sqrt{n}}\sum_{i=1}^n (R_{i,n}^2 - 2)\right)\left(\frac{1}{n}\sum_{i=1}^n R_{i,n}^2 + 2\right) \\ &= \left(\frac{1}{\sqrt{n}}\sum_{i=1}^n (R_{i,n}^2 - 2)\right)(4 + o_P(1)) \\ &= \frac{4}{\sqrt{n}}\sum_{i=1}^n (R_{i,n}^2 - 2) + o_P(1). \end{aligned}$$

■

Es gilt

$$\begin{aligned} E\left(\frac{1}{n}\sum_{i=1}^n (R_{i,n}^4 - 8)\right)^2 &= \frac{1}{n}E(R_{1,n}^4 - ER_{1,n}^4)^2 + \frac{1}{n}(ER_{1,n}^4 - 8)^2 \\ &\leq \frac{1}{n}ER_{1,n}^8 + \frac{1}{n}(ER_{1,n}^4 - 8)^2 \\ &\leq \frac{1}{n}840(1 + \kappa_n^4) + \frac{1}{n}(2\kappa_n + 3\kappa_n^2 - 5)^2, \end{aligned}$$

wobei die letzte Ungleichung aus (3.41) folgt. Aus $\kappa_n \rightarrow 1$ ergibt sich das folgende Gesetz der großen Zahlen

$$\frac{1}{n}\sum_{i=1}^n R_{i,n}^4 = 8 + o_P(1).$$

Wie in (3.43) lässt sich die Beschränktheit von

$$E\left(\frac{1}{\sqrt{n}}\sum_{i=1}^n (R_{i,n}^4 - 8)\right)^2$$

und damit die stochastische Beschränktheit von $\frac{1}{\sqrt{n}} \sum_{i=1}^n (R_{i,n}^4 - 8)$ nachweisen. Insgesamt erhalten wir also

$$\begin{aligned} & \sqrt{n} \left(T_n - \frac{1}{2} \right) \\ &= \frac{1}{2\sqrt{n}} \sum_{i=1}^n [(R_{i,n}^2 - 2) - \frac{1}{8}(R_{i,n}^4 - 8)] + o_P(1) \\ &= \frac{1}{2\sqrt{n}} \sum_{i=1}^n [(R_{i,n}^2 - \frac{1}{8}R_{i,n}^4 - 1)] + o_P(1). \end{aligned} \quad (3.44)$$

Es gilt

$$\begin{aligned} \sqrt{n}a_n &: = \sqrt{n}E[(R_{i,n}^2 - \frac{1}{8}R_{i,n}^4 - 1)] \\ &= \sqrt{n}[1 + \kappa_n - \frac{1}{8}(E(U_i^2 + \kappa_n V_i^2)^2 - 1)] \\ &= \sqrt{n}[2 + h/\sqrt{n} - \frac{1}{8}(3 + 3\kappa_n^2 + 2\kappa_n) - 1] + o(1) \\ &= \sqrt{n}[2 + h/\sqrt{n} - \frac{1}{8}(3 + 3(1 + h/\sqrt{n})^2 + 2(1 + h/\sqrt{n})) - 1] + o(1) \\ &= 0 + o(1) = o(1) \end{aligned}$$

und somit

$$\sqrt{n}a_n = o(1). \quad (3.45)$$

Satz 3.23 *Unter den Voraussetzungen (3.39) und (3.42) gilt*

$$\mathcal{L} \left(\sqrt{n} \left(T_n - \frac{1}{2} \right) \right) \Rightarrow N(0, \frac{1}{4}).$$

Beweis. Wir setzen

$$Z_{n,i} = \frac{1}{2}(R_{i,n}^2 - \frac{1}{8}R_{i,n}^4 - 1) - a_n.$$

Wegen (3.44) reicht es aus,

$$\mathcal{L} \left(\frac{1}{\sqrt{n}} \sum_{i=1}^n Z_{i,n} \right) \Rightarrow N(0, \frac{1}{4})$$

nachzuweisen. Unter den Voraussetzungen (3.39) und (3.42) erhalten wir aus (3.41)

$$\sup_{i,n} EZ_{i,n}^4 < \infty.$$

Für die Anwendung des zentralen Grenzwertsatzes, siehe Lemma 5.11, muss nur noch die Varianz von $Z_{i,n}$ untersucht werden, die nur von n und nicht von i abhängt. Weil $Z_{i,n}^2$ ein Polynom in κ_n ist und $\kappa_n \rightarrow 1$ gilt, erhalten wir

$$\begin{aligned} \lim_{n \rightarrow \infty} V(Z_{i,n}) &= \lim_{n \rightarrow \infty} EZ_{i,n}^2 = E(\lim_{n \rightarrow \infty} Z_{i,n}^2) \\ &= \frac{1}{4}E \left((U^2 + V^2) - \frac{1}{8}(U^2 + V^2)^2 - 1 \right)^2, \end{aligned}$$

wobei U und V standardnormal und unabhängig sind. Z habe eine Standardexponentialverteilung. Ausquadrieren der Klammer und die Verwendung von (3.40) ergeben

$$\begin{aligned}\lim_{n \rightarrow \infty} V(Z_{i,n}) &= \frac{1}{4} E \left(2Z - \frac{1}{2} Z^2 - 1 \right)^2 \\ &= \frac{1}{4} E \left(\frac{1}{4} Z^4 - 2Z^3 + 5Z^2 - 4Z + 1 \right) \\ &= \frac{1}{4} \left(\frac{1}{4} \cdot 4! - 2 \cdot 3! + 5 \cdot 2! - 4 + 1 \right) = \frac{1}{4}.\end{aligned}$$

■

Mit Hilfe des oben bewiesenen Grenzwertsatzes wollen wir einen asymptotischen α -Test konstruieren. Weil die Statistik T_n bei fester Alternative für $n \rightarrow \infty$ gegen $g(\kappa) < 1/2$ strebt, konvergiert die Folge der Teststatistiken W_n gegen $-\infty$. Deshalb lehnen wir bei der Konstruktion des asymptotischen Tests die Nullhypothese für kleine Werte von W_n ab. Sei z_α das α -Quantil der Standardnormalverteilung. Wir setzen

$$\varphi_n = \begin{cases} 1 & \text{falls } W_n \leq z_\alpha \\ 0 & \text{falls } W_n > z_\alpha \end{cases}.$$

Satz 3.24 Seien $(X_1, Y_1), \dots, (X_n, Y_n)$ unabhängig mit der Verteilung N_{0, Σ_i} und hat Σ_i die von i unabhängigen Eigenwerte λ_1, λ_2 . Unter der Nullhypothese $H_0 : \lambda_1 = \lambda_2 = \lambda$ ist die Folge von Tests $\{\varphi_n\}$ ein asymptotischer α -Test, d. h. es gilt

$$\lim_{n \rightarrow \infty} E_{\lambda, \lambda} \varphi_n = \alpha.$$

Die Folge von Tests $\{\varphi_n\}$ ist global konsistent, d. h. für $\lambda_1 \neq \lambda_2$ gilt

$$\lim_{n \rightarrow \infty} E_{\lambda_1, \lambda_2} \varphi_n = 1.$$

Beweis. Die erste Behauptung folgt aus der Konstruktion von φ_n und Satz 3.23. Gilt jetzt $\lambda_1 \neq \lambda_2$, so konvergiert T_n stochastisch gegen einen Wert $g(\kappa)$, der kleiner als $g(1)$ ist, was auf Grund der Konstruktion von φ_n die Konsistenz der Testfolge nach sich zieht. ■

Bemerkung 3.25 Dem Grenzwertsatz 3.23 können wir entnehmen, dass unter der lokalen Alternative (3.42) die asymptotische Güte des Tests φ_n gerade das Niveau α ist. Lokale Alternativen mit der Geschwindigkeit $1/\sqrt{n}$ werden von diesem Test nicht erkannt. Unabhängig davon ist er aber global konsistent, wie oben bewiesen wurde.

3.4 Bedingte Tests auf Vorliegen einer Hauptrichtung

Die Schwierigkeit der asymptotischen Tests besteht darin, dass die Fehlerwahrscheinlichkeit erster Art nicht exakt eingehalten wird. Bestehen unter der Nullhypothese die Daten aus zwei unabhängigen Bestandteilen, so lassen sich bedingte Tests konstruieren, die dann exakte Tests in dem Sinne sind, dass sie das Niveau exakt einhalten. Sind die Daten eindimensional, so sind solche unabhängigen Bestandteile die Rangstatistik und die Ordnungsstatistik. In unserem Falle sind solche unabhängigen Größen gerade der Winkel und die Länge.

3 Testen auf Hauptrichtungen des Wachstums

Wir wollen kurz auf das Konzept bedingter Tests eingehen. Gegeben sei ein statistisches Modell, wo der Stichprobenraum eine Produktstruktur hat

$$(\mathcal{U} \times \mathcal{V}, \mathfrak{U} \otimes \mathfrak{V}, (Q_\theta)_{\theta \in \Theta}).$$

Jetzt sei das Testproblem

$$H_0 : \theta \in \Theta_0 \quad \text{gegen} \quad H_A : \theta \in \Theta_A. \quad (3.46)$$

gegeben mit $\Theta_0 \cup \Theta_A = \Theta$ und $\Theta_0 \cap \Theta_A = \emptyset$. Für eine Funktion $S : \mathcal{U} \times \mathcal{V} \rightarrow \mathbb{R}$ soll ein α -Test konstruiert werden, der die Nullhypothese für große Werte von $S(U, V)$ ablehnt. Wir bezeichnen durch $K_\theta(\cdot|v)$ die bedingte Verteilung von $S(U, V)$ unter der Bedingung $V = v$ und durch P_θ die Verteilung von V . Wir setzen voraus, dass die bedingte Verteilung K_θ von $S(U, V)$ bei gegebenem $V = v$ für $\theta \in \Theta_0$ unabhängig von θ ist. Es gibt also einen stochastischen Kern K derart, dann für jede beschränkte und messbare Funktion f gilt

$$\int f(u, v) Q_\theta(du, dv) = \int \left(\int f(u, v) K(du|v) P_\theta(dv) \right), \quad \text{für alle } \theta \in \Theta_0. \quad (3.47)$$

Zur Konstruktion eines α -Tests wählen wir für festes v einen Test $b(\cdot, v)$ mit der Eigenschaft

$$\int b(s, v) K(ds|v) = \alpha. \quad (3.48)$$

Dann gilt

$$E_\theta b(S(U, V), V) = \int \left(\int b(s, v) K(ds|v) P_\theta(dv) \right) = \alpha, \quad \text{für alle } \theta \in \Theta_0 \quad (3.49)$$

und damit ist $b(S(u, v), v)$ ein α -Test, der als *bedingter Test* bezeichnet wird.

Wir wenden das Konzept der bedingten Tests auf homogene Verteilungen an. Wir betrachten das Modell in (3.34) mit $f = h_{g,h,k}$ und

$$h_{g,h,k,A}(x, y) := ((\det(A))^{-1} (g(A^{-1}(x, y)^T) k(A^{-1}(x, y)^T)), \quad A \in \mathfrak{S}, g \in \mathfrak{G}_{p,q},$$

$$h_{g,h,k} \text{ erfüllt (2.17).}$$

Besitzt der Vektor (X, Y) die Dichte $h_{g,h,k,A}$, dann gilt nach Folgerung 2.10

$$\begin{pmatrix} X \\ Y \end{pmatrix} = L \cdot h^{-1/p}(\Phi) \begin{pmatrix} \cos \Phi \\ \sin \Phi \end{pmatrix}, \quad (3.50)$$

$$L = R h^{1/p}(\Phi) \text{ und } R = (X^2 + Y^2)^{1/2},$$

wobei gilt

L und Φ sind unabhängig,

L^p hat die Verteilung G mit der Dichte $t^{\alpha(p,q)-1} I_{(0,\infty)}(t) g(t)$,

Φ hat die Verteilung $P_{h,k}$ mit der Dichte $\frac{k(\varphi)}{p h^{\alpha(p,q)}(\varphi)} I_{[0,2\pi)}(\varphi)$.

Liegen jetzt n unabhängige Beobachtungen $(X_1, Y_1), \dots, (X_n, Y_n)$ mit der Dichte $h_{g,h,k,A}$ vor, so haben wir folgendes statistische Modell

$$\begin{aligned} & (\mathcal{U} \times \mathcal{V}, \mathfrak{U} \otimes \mathfrak{V}, (Q_\theta)_{\theta \in \Theta}) \\ &= ([0, 2\pi)^n \times (0, \infty)^n, \mathfrak{B}_{[0,2\pi)}^{\otimes n} \otimes \mathfrak{B}_{(0,\infty)}^{\otimes n}, (P_{h,k,A}^{\otimes n} \otimes G^{\otimes n})_{A \in \mathfrak{S}}), \\ & U, V \text{ Projektionen von } \mathcal{U} \times \mathcal{V} \text{ auf } \mathcal{U} \text{ bzw. } \mathcal{V}. \end{aligned} \quad (3.51)$$

Wir betrachten jetzt ähnlich wie in (3.27), die Teststatistik

$$S = \frac{1}{(\sum_{i=1}^n (X_i^2 + Y_i^2))^2} (\sum_{i=1}^n X_i^2 - \sum_{i=1}^n Y_i^2)^2 + (2 \sum_{i=1}^n X_i Y_i)^2. \quad (3.52)$$

Für die Statistik S ergibt sich dann aus (3.50)

$$\begin{aligned} S &= S_{L_1, \dots, L_n}(\Phi_1, \dots, \Phi_n), \quad \text{mit} \\ &S_{l_1, \dots, l_n}(\varphi_1, \dots, \varphi_n) \\ &= \frac{1}{(\sum_{i=1}^n l_i^2)^2} \left(\left(\sum_{i=1}^n l_i^2 \frac{\cos^2 \varphi_i - \sin^2 \varphi_i}{h^{2/p}(\varphi_i)} \right)^2 + \left(2 \sum_{i=1}^n l_i^2 \frac{\cos \varphi_i \cdot \sin \varphi_i}{h^{2/p}(\varphi_i)} \right)^2 \right). \end{aligned}$$

Für U, V aus (3.51) setzen wir $U = (\Phi_1, \dots, \Phi_n), V = (L_1, \dots, L_n)$ und betrachten die Φ_i und L_i als Zufallsvariable auf dem Wahrscheinlichkeitsraum (3.51). Sei

$$\begin{aligned} K_{n,A}(\cdot | l_1, \dots, l_n) &= \text{die bedingte Verteilung von } S \text{ unter der Bedingung,} \\ L_1 &= l_1, \dots, L_n = l_n \text{ unter der Verteilung } P_A^{\otimes n} \otimes G^{\otimes n}. \end{aligned}$$

Wegen der Unabhängigkeit von U und V gilt

$$K_{n,A}(\cdot | l_1, \dots, l_n) = P_A^{\otimes n} \circ S_{l_1, \dots, l_n}^{-1}.$$

Wegen (2.35) in Folgerung 2.14 hängen die Verteilungen $P_A^{\otimes n}$ unter der Nullhypothese $\Sigma = \sigma^2 I$ nicht von σ^2 ab. Dann ist auch $K_{n,\sigma^2 I}(\cdot | l_1, \dots, l_n)$ unter der Nullhypothese von σ^2 unabhängig. Weiterhin ist die bedingte Verteilung $K_{n,A}$ die Verteilung von Summen unabhängiger Zufallsvariablen, deren Verteilungsfunktionen stetig sind. Dann sind die Verteilungsfunktionen

$$F_n(t | l_1, \dots, l_n) = K_{n,I}((0, t] | l_1, \dots, l_n)$$

wieder stetig. Durch $c_n(\gamma | l_1, \dots, l_n)$ bezeichnen wir die entsprechenden γ -Quantile mit deren Hilfe wir die bedingten Tests konstruieren. Hierzu setzen wir

$$b_n(s, l_1, \dots, l_n) = \begin{cases} 1 & \text{falls } S(\varphi_1, \dots, \varphi_n, l_1, \dots, l_n) > c_n(1 - \alpha | l_1, \dots, l_n) \\ 0 & \text{sonst} \end{cases}.$$

\mathfrak{P}_{abs} bezeichne die Menge der absolutstetigen Verteilungen auf $(0, \infty)$. Dann liegt folgendes Testproblem vor

$$\begin{aligned} H_0 &: P_A^{\otimes n} \otimes G^{\otimes n}, AA^T \in \mathfrak{S}_0, G \in \mathfrak{P}_{abs} \text{ gegen} \\ H_A &: P_\Sigma^{\otimes n} \otimes G^{\otimes n}, \Sigma \in \mathfrak{S}_A, G \in \mathfrak{P}_{abs}, \\ &G \text{ ist unbekannt.} \end{aligned} \quad (3.53)$$

Aus der Konstruktion der Tests b_n sowie der Beziehung (3.49) ergibt sich folgende Aussage.

Satz 3.26 Die Tests b_n sind bedingte α -Tests für das Modell (3.51) und das Testproblem (3.53).

Bei der Anwendung der bedingten Tests benötigt man die Quantile c_n , die jeweils noch von der entsprechenden Bedingung abhängen. Weil sich die bedingten Verteilungen K_n nicht geschlossen berechnen lassen, muss man die Quantile durch Simulationen oder durch geeignete Grenzwertsätze approximieren.

3.5 Computersimulationen

3.5.1 Niveau der Tests

Monte-Carlo-Simulationen sind ein systematisches Hilfsmittel, um zu überprüfen, ob asymptotische α -Tests das Niveau bereits bei kleinen oder mittleren Stichproben einhalten. Gleichzeitig geben solche Simulationen Einblicke in die Struktur der Gütefunktion für fest gewählte Alternativen.

In diesem Abschnitt wollen wir den in Abschnitt 3.2 konstruierten asymptotischen Test und den in Abschnitt 3.4 konstruierten bedingten Test auf bivariate elliptisch konturierte Verteilungen anwenden. Deren Dichten wurden in (2.15) definiert und haben die folgende Form

$$\varphi_{g,\Sigma}(x, y) = \left(\pi \sqrt{\det(\Sigma)} \right)^{-1} g((x, y) \Sigma^{-1} (x, y)).$$

Σ ist die Kovarianzmatrix von (X, Y) , sie ist symmetrisch und positiv definit. Die elliptisch konturierten Verteilungen unterscheiden sich durch die die Dichte generierende Funktion g . Für die Simulation betrachten wir drei Verteilungsklassen mit folgenden g_i für $i = 1, 2$ und 3 .

- Normalverteilung: In (2.36) ist

$$g_1(t) = I_{(0,\infty)} \frac{1}{2} \exp \left\{ -\frac{t}{2} \right\}.$$

- Kotz-Verteilung: In (2.36) ist

$$g_{\alpha,\beta,\gamma}(t) = C(\alpha,\beta,\gamma) I_{(0,\infty)}(t) t^{\alpha-1} \exp\{-\beta t^\gamma\}, \quad \alpha, \beta, \gamma > 0.$$

Für die Simulation betrachten wir einen Spezialfall von $g_{\alpha,\beta,\gamma}(t)$ mit $\alpha = \frac{1}{2}$, $\beta = \gamma = 1$, d. h.

$$g_2(t) = I_{(0,\infty)} \frac{1}{\Gamma\left(\frac{1}{2}\right)} t^{-\frac{1}{2}} \exp\{-t\}.$$

- Pearson-Typ-VII Verteilung: In (2.36) ist

$$g_{\alpha,\beta}(t) = C(\alpha,\beta) I_{(0,\infty)}(1 + \beta t)^{-(1+\alpha)}, \quad \alpha \geq \frac{3}{2}, \beta > 0.$$

Für die Simulation betrachten wir einen Spezialfall von $g_{\alpha,\beta}(t)$ mit $\alpha = \frac{3}{2}$ und $\beta = \frac{1}{3}$, d. h.

$$g_3(t) = I_{(0,\infty)} \frac{1}{2} \left(1 + \frac{t}{3}\right)^{-(1+\frac{3}{2})}.$$

Obwohl für normalverteilte Daten der Mauchly-Test für jeden Stichprobenumfang die Fehlerwahrscheinlichkeit erster Art einhält, führen wir trotzdem Simulationen durch, um für die anderen Tests mögliche Abweichungen des geschätzten Niveaus vom vorgegebenen Niveau α besser bewerten zu können.

Wir betrachten das Testproblem in (3.34) und verwenden die Teststatistiken in (3.30) und (3.52). Die Simulationen wurden mit Hilfe des Softwarepaketes R ausgeführt.

1. Für $n = 20, 50$ und 100 erzeugen wir $(X_1, Y_1), \dots, (X_n, Y_n)$ mit der Dichtefunktion $\varphi_{g_i,\Sigma}(x, y)$ mit $\Sigma = I$, $i = 1, 2$ bzw. 3 .

3 Testen auf Hauptrichtungen des Wachstums

2. Wir berechnen die Werte der Teststatistiken \hat{Q}_n in (3.30) und S in (3.52).

3. Wir führen die Tests

$$\psi_n = \begin{cases} 1 & , \quad \hat{Q}_n > -2 \ln(\alpha) \\ 0 & , \quad \text{sonst} \end{cases} \quad \text{und} \quad b_n = \begin{cases} 1 & , \quad S > c_n(1 - \alpha | t_1, \dots, t_n) \\ 0 & , \quad \text{sonst} \end{cases}$$

durch, wobei $c_n(1 - \alpha | t_1, \dots, t_n)$ der durch Simulation geschätzte Quantilwert des bedingten Tests ist.

4. Die Schritte 1.-3. werden N -mal wiederholt und das tatsächliche Konfidenzniveau wird geschätzt durch

$$\hat{\alpha} = \frac{\text{Anzahl der Ablehnungen von } H_0}{N}.$$

In unseren Simulationen haben wir $N = 1000$ verwendet.

In der Tabelle 3.2 bezeichnen $\alpha_{Asp.}$ und $\alpha_{Bed.}$ die geschätzten Niveaus des asymptotischen Tests und des bedingten Tests.

Normalverteilung			
α	$n = 20$	$n = 50$	$n = 100$
0.01	$\alpha_{Aymp.} = 0.012$	$\alpha_{Aymp.} = 0.009$	$\alpha_{Aymp.} = 0.006$
	$\alpha_{Bed.} = 0.007$	$\alpha_{Bed.} = 0.007$	$\alpha_{Bed.} = 0.010$
0.05	$\alpha_{Aymp.} = 0.049$	$\alpha_{Aymp.} = 0.049$	$\alpha_{Aymp.} = 0.056$
	$\alpha_{Bed.} = 0.042$	$\alpha_{Bed.} = 0.049$	$\alpha_{Bed.} = 0.049$
0.1	$\alpha_{Aymp.} = 0.094$	$\alpha_{Aymp.} = 0.11$	$\alpha_{Aymp.} = 0.103$
	$\alpha_{Bed.} = 0.088$	$\alpha_{Bed.} = 0.091$	$\alpha_{Bed.} = 0.099$
Kotz-Verteilung			
0.01	$\alpha_{Aymp.} = 0.012$	$\alpha_{Aymp.} = 0.012$	$\alpha_{Aymp.} = 0.011$
	$\alpha_{Bed.} = 0.008$	$\alpha_{Bed.} = 0.009$	$\alpha_{Bed.} = 0.011$
0.05	$\alpha_{Aymp.} = 0.069$	$\alpha_{Aymp.} = 0.054$	$\alpha_{Aymp.} = 0.049$
	$\alpha_{Bed.} = 0.045$	$\alpha_{Bed.} = 0.052$	$\alpha_{Bed.} = 0.050$
0.1	$\alpha_{Aymp.} = 0.140$	$\alpha_{Aymp.} = 0.113$	$\alpha_{Aymp.} = 0.101$
	$\alpha_{Bed.} = 0.097$	$\alpha_{Bed.} = 0.100$	$\alpha_{Bed.} = 0.097$
Pearson-Typ-VII-Verteilung			
0.01	$\alpha_{Aymp.} = 0$	$\alpha_{Aymp.} = 0.0001$	$\alpha_{Aymp.} = 0.01$
	$\alpha_{Bed.} = 0.009$	$\alpha_{Bed.} = 0.013$	$\alpha_{Bed.} = 0.010$
0.05	$\alpha_{Aymp.} = 0$	$\alpha_{Aymp.} = 0.005$	$\alpha_{Aymp.} = 0.056$
	$\alpha_{Bed.} = 0.048$	$\alpha_{Bed.} = 0.054$	$\alpha_{Bed.} = 0.050$
0.1	$\alpha_{Aymp.} = 0.004$	$\alpha_{Aymp.} = 0.025$	$\alpha_{Aymp.} = 0.106$
	$\alpha_{Bed.} = 0.094$	$\alpha_{Bed.} = 0.100$	$\alpha_{Bed.} = 0.097$

Tabelle 3.2: Geschätzte Niveaus des asymptotischen Tests $\alpha_{Asp.}$ und des bedingten Tests $\alpha_{Bed.}$

In der Tabelle 3.2 können wir ablesen, dass die Fehlerwahrscheinlichkeit erster Art des asymptotischen Tests für kleine n nicht exakt eingehalten wurde, besonders für die Kotz-verteilten Stichproben. Bei dem bedingten Test wurde sie deutlich besser eingehalten.

3.5.2 Die asymptotische Güte unter speziellen Alternativen

Für die Untersuchung der Güte obiger Tests betrachten wir die Tests in normalverteilten, Kotz-verteilten bzw. Pearson-Typ-VII-verteilten Grundgesamtheiten. Wir betrachten zwei feste Alternativen:

1. Inhomogene Kovarianzmatrix: $\Sigma_A = \begin{pmatrix} \sigma^2 & 0 \\ 0 & \tau^2 \end{pmatrix}$ mit $\sigma^2 \neq \tau^2$, $\sigma^2, \tau^2 \in (0, \infty)$ und damit sind $\lambda_1 = \sigma^2$ und $\lambda_2 = \tau^2$. Für die Simulation setzen wir $\sigma^2 = 1 + \gamma$ und $\tau^2 = 1 - \gamma$ mit $\gamma \in (0, 1)$.
2. Korrelierte Kovarianzmatrix: $\Sigma_A = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$ mit $\rho \neq 0$ und damit sind $\lambda_1 = 1 + \rho$ und $\lambda_2 = 1 - \rho$.

Wir erzeugen zufällige Punkte mit den zweidimensionalen Dichtefunktionen $\varphi_{g_i, \Sigma_A}(x, y)$, $i = 1, 2$ bzw. 3 und benutzen dazu Satz 2.18 in Kapitel 2. Dieser Satz gestattet es, unabhängige Zufallsgrößen L^2 und Φ zu simulieren mit den in Satz 2.18 angegebenen Dichten

$$\begin{aligned} f_{L^2}(t) &= l_{(0, \infty)}(t) g_i(t), \quad i = 1, 2 \text{ bzw. } 3, \\ f_{\Phi}(\varphi) &= \frac{1}{2\pi} l_{[0, 2\pi)}(\varphi) \frac{1}{\sqrt{\lambda_1 \lambda_2 \rho_{\Sigma_A}(\varphi)}}. \end{aligned}$$

Die zufälligen Punkte $(X_1, Y_1), \dots, (X_n, Y_n)$ sind dann Realisierungen von (X, Y) und entstehen durch Anwendung der folgenden Beziehung

$$\begin{pmatrix} X \\ Y \end{pmatrix} = L \begin{pmatrix} \cos \Phi / \sqrt{\rho_{\Sigma_A}(\Phi)} \\ \sin \Phi / \sqrt{\rho_{\Sigma_A}(\Phi)} \end{pmatrix}.$$

Die Abbildungen 3.3, 3.4 und 3.5 sowie 3.6, 3.7 und 3.8 geben für $n = 20, 50$ und 100 sowie $\alpha = 0.05$ den Verlauf der Gütefunktionen der beiden Tests an. Die rote Linie zeigt jeweils die Gütefunktion des asymptotischen Tests, die blaue die des bedingten Tests.

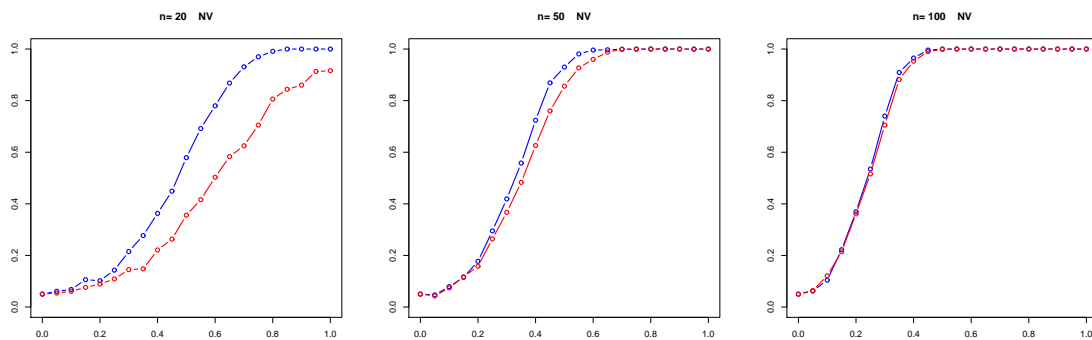


Abbildung 3.3: Verlauf der Gütefunktion des asymptotischen Tests bzw. des bedingten Tests in normalverteilten Grundgesamtheiten mit $n = 20, 50$ bzw. 100 unter der ersten Alternative

3 Testen auf Hauptrichtungen des Wachstums

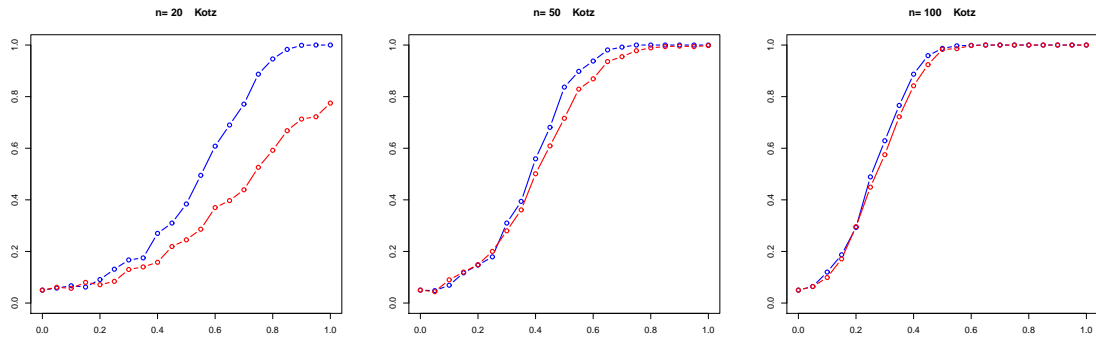


Abbildung 3.4: Verlauf der Gütefunktion des asymptotischen Tests bzw. des bedingten Tests in Kotz-verteilten Grundgesamtheiten mit $n = 20, 50$ bzw. 100 unter der ersten Alternative

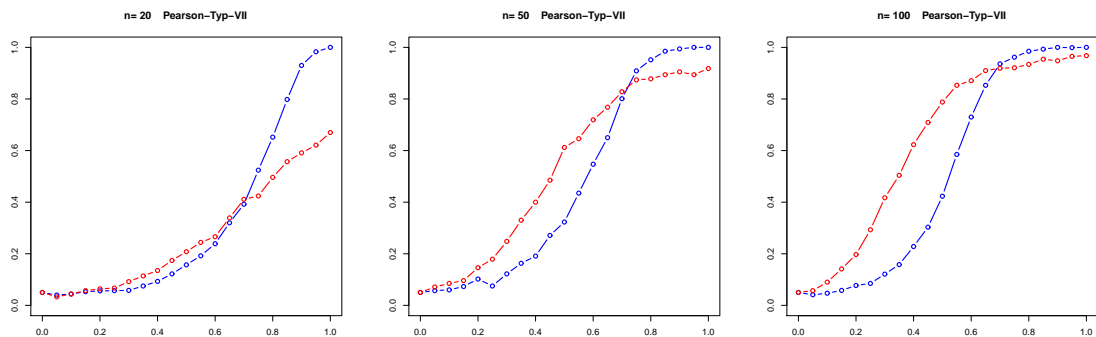


Abbildung 3.5: Verlauf der Gütefunktion des asymptotischen Tests bzw. des bedingten Tests in Pearson-Typ-VII-verteilten Grundgesamtheiten mit $n = 20, 50$ bzw. 100 unter der ersten Alternative

3.6 Test von Projektdaten

In diesem Abschnitt wollen die Tests auf Projektdaten anwenden, um zu sehen, ob eine Hauptwachstumsrichtung existiert, in der es ein besonders starkes Wachstum der Filamente gibt. Die Daten zum Testen wurden von Herrn Birkholz gewonnen. Die Daten messen die Längen und die Winkel der Aktinfilamente von Osteoblasten auf drei Oberflächen: *Daten 1* und *Daten 2* auf der blanken Titanoberfläche, *Daten 3* auf der gräbenstrukturierten Titanoberfläche und *Daten 4*, *Daten 5* sowie *Daten 6* auf der pfostenstrukturierten Titanoberfläche. Die zugehörigen Datensätze sind auf der beigefügten CD als CSV-Dateien zu finden.

Ein Datensatz der Testdaten besteht aus 180 Paaren (L_i, W_i) von Längen und Winkeln der Aktinfilamente. Die Bilder in Abbildung 3.9 stellen die Beziehung von Längen und Winkeln dar. Auf manchen Oberflächen haben die Filamente kein ausgeprägtes Wachstum in eine Richtung. Ein Beispiel dafür sieht man im Bild *Daten 1* in Abbildung 3.9. Auf anderen Oberfläche zeigen die Filamente ein ausgeprägtes Wachstum in eine Richtung. Dieser Effekt wird in dem Bild *Daten 3* in Abbildung 3.9 illustriert.

Wir testen im Folgenden auf die Existenz einer Hauptrichtung mit dem asymptotischen Test aus Abschnitt 3.2 und dem bedingten Test aus Abschnitt 3.4. Wir betrachten das Testproblem

3 Testen auf Hauptrichtungen des Wachstums

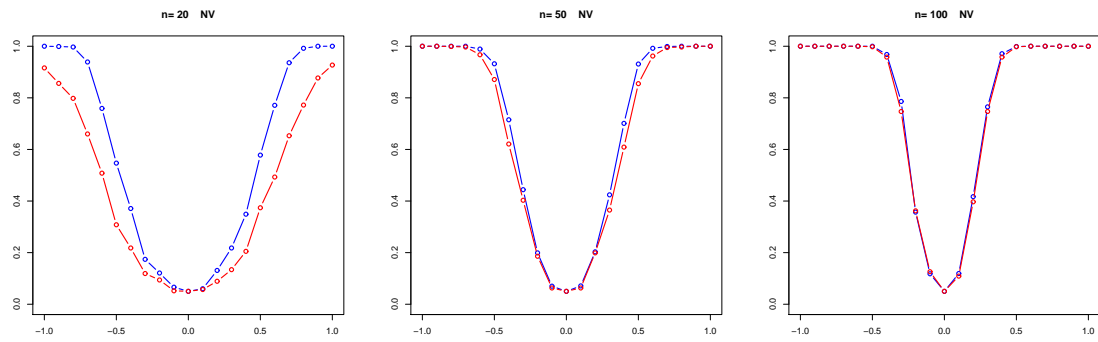


Abbildung 3.6: Verlauf der Gütefunktion des asymptotischen Tests bzw. des bedingten Tests in normalverteilten Grundgesamtheiten mit $n = 20, 50$ bzw. 100 unter der zweiten Alternative

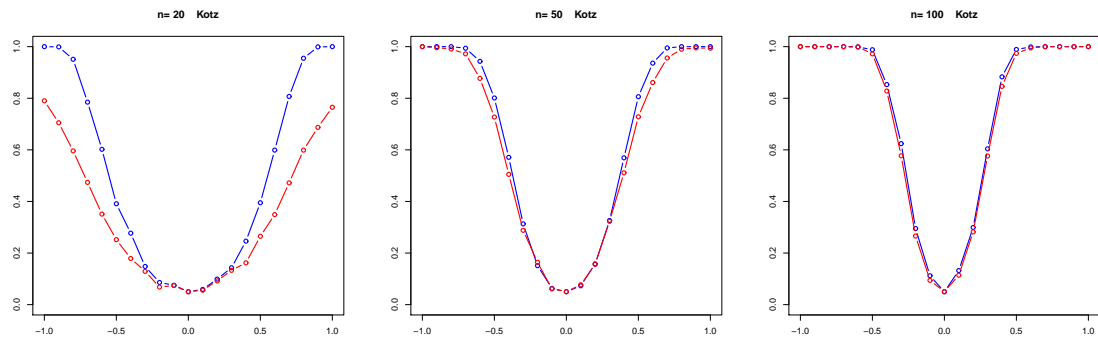


Abbildung 3.7: Verlauf der Gütefunktion des asymptotischen Tests bzw. des bedingten Tests in Kotz-verteilten Grundgesamtheiten mit $n = 20, 50$ bzw. 100 unter der zweiten Alternative

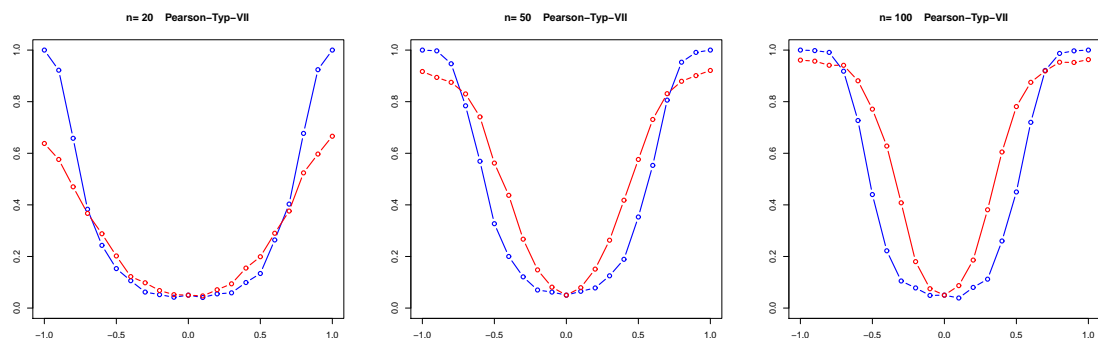


Abbildung 3.8: Verlauf der Gütefunktion des asymptotischen Tests bzw. des bedingten Tests in Pearson-Typ-VII-verteilten Grundgesamtheiten mit $n = 20, 50$ bzw. 100 unter der zweiten Alternative

3 Testen auf Hauptrichtungen des Wachstums

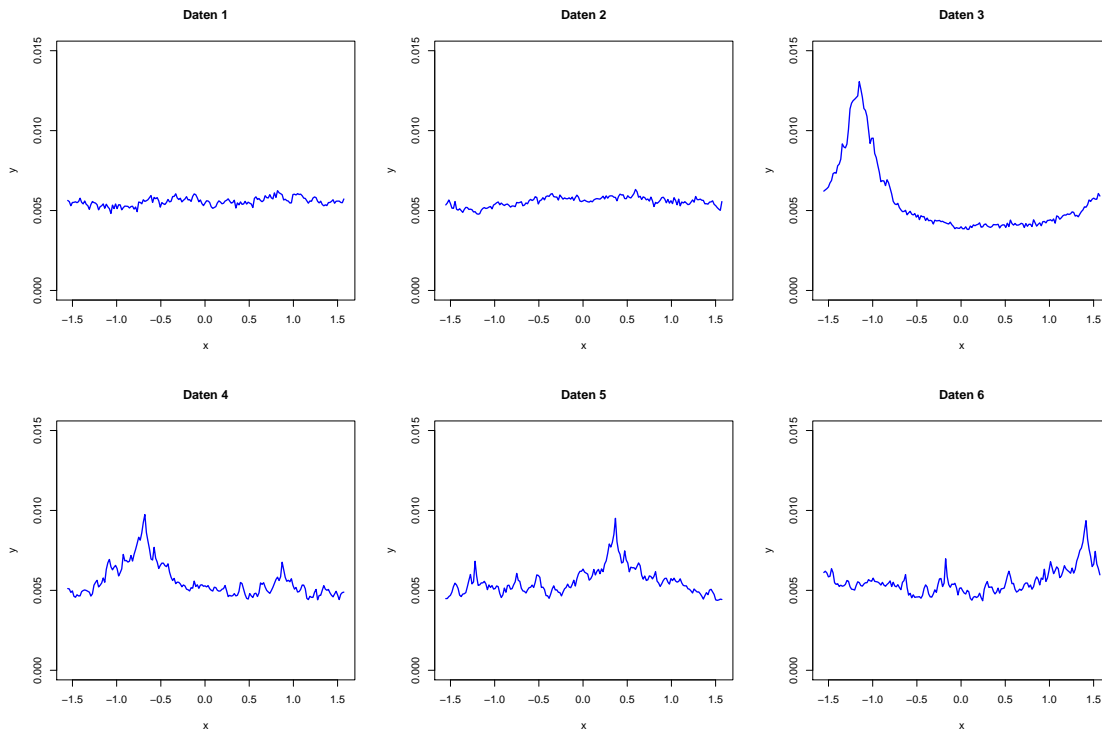


Abbildung 3.9: Kurven der Längen der Filamente bei gegebenen Winkeln

in (3.34) und verwenden die Teststatistiken in (3.30) und (3.52). In der Tabelle 3.3 sind die Testergebnisse des asymptotischen Tests dargestellt und die Tabelle 3.4 zeigt die Testergebnisse des bedingten Tests.

Bei der Anwendung des asymptotischen Tests wird die Nullhypothese abgelehnt, obwohl die Bilder 1 und 2 in der Abbildung 3.9 keine klare Hauptrichtung ausweisen. Dieser unerwünschte Effekt könnte dadurch zustande kommen, dass diese Daten insgesamt vom Modell abweichen. Beim bedingten Test erfolgt in den Datensätzen 1, 2 und 6 keine Ablehnung in Übereinstimmung mit den Bildern. Vermutlich ist also der bedingte Test robuster gegenüber Modellabweichungen. Weil er außerdem ein exakter α -Test ist, ist er gegenüber dem asymptotischen Test vorzuziehen.

3 Testen auf Hauptrichtungen des Wachstums

			Kritische Werte	
		$\alpha = 0.1$	$\alpha = 0.05$	$\alpha = 0.01$
	Werte der Teststatistik	$c_{n,1-\alpha} = 4.6052$	$c_{n,1-\alpha} = 5.9915$	$c_{n,1-\alpha} = 9.2103$
Daten 1	57.8314	Ablehnen	Ablehnen	Ablehnen
Daten 2	49.1491	Ablehnen	Ablehnen	Ablehnen
Daten 3	102.3382	Ablehnen	Ablehnen	Ablehnen
Daten 4	74.4140	Ablehnen	Ablehnen	Ablehnen
Daten 5	41.0842	Ablehnen	Ablehnen	Ablehnen
Daten 6	74.0667	Ablehnen	Ablehnen	Ablehnen

Tabelle 3.3: Testergebnisse des asymptotischen Tests

			Kritische Werte	
	Werte der Teststatistik	$\alpha = 0.1$	$\alpha = 0.05$	$\alpha = 0.01$
Quantile 1		$c_{n,1-\alpha} = 0.0136$	$c_{n,1-\alpha} = 0.0165$	$c_{n,1-\alpha} = 0.0247$
Daten 1	0.0044	<i>Annehmen</i>	<i>Annehmen</i>	<i>Annehmen</i>
Quantile 2		$c_{n,1-\alpha} = 0.0130$	$c_{n,1-\alpha} = 0.0169$	$c_{n,1-\alpha} = 0.0268$
Daten 2	0.0033	<i>Annehmen</i>	<i>Annehmen</i>	<i>Annehmen</i>
Quantile 3		$c_{n,1-\alpha} = 0.0240$	$c_{n,1-\alpha} = 0.0318$	$c_{n,1-\alpha} = 0.0479$
Daten 3	0.2139	Ablehnen	Ablehnen	Ablehnen
Quantile 4		$c_{n,1-\alpha} = 0.0151$	$c_{n,1-\alpha} = 0.0197$	$c_{n,1-\alpha} = 0.0301$
Daten 4	0.0310	Ablehnen	Ablehnen	Ablehnen
Quantile 5		$c_{n,1-\alpha} = 0.0142$	$c_{n,1-\alpha} = 0.0182$	$c_{n,1-\alpha} = 0.0265$
Daten 5	0.0209	Ablehnen	Ablehnen	<i>Annehmen</i>
Quantile 6		$c_{n,1-\alpha} = 0.0147$	$c_{n,1-\alpha} = 0.0184$	$c_{n,1-\alpha} = 0.0283$
Daten 6	0.0138	<i>Annehmen</i>	<i>Annehmen</i>	<i>Annehmen</i>

Tabelle 3.4: Testergebnisse des bedingten Tests

4 Anpassungstests

4.1 Grenzwertsätze für die empirische Verteilungsfunktion

4.1.1 Konvergenz unter einfacher Nullhypothese

Es seien X_1, X_2, \dots unabhängige und identisch verteilte Zufallsvariablen mit der Verteilungsfunktion F . Es seien

$$\begin{aligned}\hat{F}_n(t) &= \frac{1}{n} \sum_{i=1}^n I_{(-\infty, t]}(X_i), \\ B_n(t) &= \sqrt{n}(\hat{F}_n(t) - F(t)).\end{aligned}$$

Im Weiteren bezeichnen wir durch \Rightarrow die schwache Konvergenz von Verteilungen, die auf dem Skorokhodraum $\mathbb{D}[-\infty, \infty]$ definiert sind. Der folgende Satz ist ein Spezialfall von Satz 11, Seite 97 in Pollard (1984).

Satz 4.1 *Ist die Verteilungsfunktion F stetig, dann gilt*

$$\mathcal{L}(B_n) \Rightarrow \mathcal{L}(B(F)),$$

wobei B die Brownsche Brücke auf $[0, 1]$ ist.

Die oben formulierte Aussage lässt sich wie folgt anwenden. Sei

$$\kappa(x) = \sup_{-\infty \leq t \leq \infty} |x(t)|. \quad (4.1)$$

Wir bezeichnen κ als Kolmogoroff-Smirnov-Funktional. κ ist offensichtlich stetig. Sei \mathbb{Q} die Menge der rationalen Zahlen. Dann gilt wegen der Rechtsstetigkeit der Funktionen x aus \mathbb{D}

$$\sup_{-\infty \leq t \leq \infty} |x(t)| = \sup_{t \in \mathbb{Q}} |x(t)|,$$

deshalb ist $\kappa : \mathbb{D}[-\infty, \infty] \rightarrow \mathbb{R}$ messbar bezüglich $\mathfrak{B}_{\mathbb{D}}$ (für die Definition siehe Anhang II). Ist B die Brownsche Brücke auf $[0, 1]$, so gilt nach Billingsley (1986)

$$P(\sup_{0 \leq t \leq 1} |B(t)| \leq x) = 1 - 2 \sum_{k=1}^{\infty} (-1)^{k+1} \exp\{-2k^2 x^2\}.$$

Andererseits gilt wegen der Stetigkeit von F

$$\sup_{-\infty \leq t \leq \infty} |B(F(t))| = \sup_{0 \leq s \leq 1} |B(s)|.$$

Aus Satz 4.1 erhalten wir deshalb folgende Aussage.

Folgerung 4.2 *Ist die Verteilungsfunktion F stetig, so gilt für $B_n = \sqrt{n}(\hat{F}_n - F)$ und alle $x > 0$*

$$\lim_{n \rightarrow \infty} P(\sup_{-\infty \leq t \leq \infty} |B_n(t)| \leq x) = 1 - 2 \sum_{k=1}^{\infty} (-1)^{k+1} \exp\{-2k^2 x^2\}.$$

4 Anpassungstests

Wir untersuchen jetzt die Statistik $\int B_n^2(t) d\hat{F}_n(t)$ näher und bezeichnen dazu durch $X_{n:1} \leq \dots \leq X_{n:n}$ die zur Stichprobe X_1, \dots, X_n gehörige Ordnungsstatistik. Wegen der Stetigkeit von F sind die X_i mit Wahrscheinlichkeit Eins alle verschieden und es gilt

$$\hat{F}_n(X_{n:i}) = \frac{i}{n}.$$

Hieraus folgt

$$\int B_n^2(t) d\hat{F}_n(t) = \sum_{i=1}^n \left(\frac{i}{n} - F(X_{n:i}) \right)^2. \quad (4.2)$$

Ist F stetig und P die zugehörige Verteilung, so gilt

$$\int_{-\infty}^{\infty} B^2(F(t)) P(dt) = \int_0^1 B^2(s) ds.$$

Aus Satz 4.1 und dem Lemma 5.8 im Anhang II ergibt sich folgende Aussage.

Proposition 4.3 *Ist die Verteilungsfunktion F stetig, so gilt für alle x*

$$\lim_{n \rightarrow \infty} P\left(\sum_{i=1}^n \left(\frac{i}{n} - F(X_{n:i})\right)^2 \leq x\right) = P\left(\int_0^1 B^2(s) ds \leq x\right).$$

4.1.2 Konvergenz unter lokalen Alternativen

Es seien P_n und P Verteilungen auf $(\mathcal{X}, \mathfrak{A})$, die der Bedingung $P_n \ll P$ genügen. Wir setzen $g_n = dP_n/dP$ und

$$a_n = 2\sqrt{n}(\sqrt{g_n} - 1). \quad (4.3)$$

Wir nehmen an, dass für ein $a \in L_2(P)$ gilt

$$\lim_{n \rightarrow \infty} \int (a_n(x) - a(x))^2 P(dx) = 0. \quad (4.4)$$

Durch $X_{n,1}, \dots, X_{n,n}$ bezeichnen wir die Koordinatenprojektionen von \mathcal{X}^n auf \mathcal{X} . Aus Lemma 5.1 im Anhang I erhalten wir für

$$L_n = \frac{dP_n^{\otimes n}}{dP^{\otimes n}}(X_{n,1}, \dots, X_{n,n})$$

folgende Entwicklung, die als LAN-Bedingung bezeichnet wird. Es gilt

$$\ln L_n = \frac{1}{\sqrt{n}} \sum_{i=1}^n a(X_{n,i}) - \frac{1}{2} \int_0^1 a^2(x) P(dx) + o_{P^{\otimes n}}(1). \quad (4.5)$$

Wir untersuchen das asymptotische Verhalten der Statistiken

$$T_n = \frac{1}{\sqrt{n}} \sum_{i=1}^n b_n(X_{n,i}),$$

wobei $\mathcal{L}(X_{n,i}) = P_n$, $i = 1, \dots, n$, $b_n, b \in L_2^0(P)$ und

$$\lim_{n \rightarrow \infty} \int (b_n - b)^2 dP = 0 \quad (4.6)$$

4 Anpassungstests

vorausgesetzt werden. Aus der Proposition 5.5 im Anhang I ergibt sich unter den Voraussetzungen (4.4) und (4.6)

$$\mathcal{L}(T_n|P_n^{\otimes n}) \Rightarrow \mathbf{N}\left(\int abdP, \int b^2 dP\right). \quad (4.7)$$

Wir untersuchen jetzt den empirischen Prozess B_n , wenn die Beobachtungen aus lokalen Alternativen kommen, d. h. wir betrachten die Folge der Wahrscheinlichkeitsräume $(\mathbb{R}_n, \mathfrak{B}^{\otimes n}, P_n^{\otimes n})$, wobei die $X_{n,i}$, $i = 1, \dots, n$ die Koordinatenabbildungen sind. Für festes $x \in \mathbb{R}$, reelle Zahlen $\alpha_1, \dots, \alpha_m$ und $t_1 < t_2 < \dots < t_m$ sei

$$b(x) = \sum_{j=1}^m \alpha_j (I_{(-\infty, t_j]}(x) - F(t_j)).$$

Dann gilt $b \in L_2^0(P)$ und mit

$$T_n = \sum_{j=1}^m \alpha_j B_n(t_j) = \frac{1}{\sqrt{n}} \sum_{i=1}^n b(X_{n,i})$$

folgt aus (4.7)

$$\mathcal{L}\left(\sum_{j=1}^m \alpha_j B_n(t_j) | P_n^{\otimes n}\right) \Rightarrow \mathbf{N}\left(\int abdP, \int b^2 dP\right).$$

Hierbei gilt

$$\int abdP = \sum_{j=1}^m \alpha_j (A(t_j) - F(t_j)),$$

wobei

$$A(t) = \int I_{(-\infty, t]}(s) a(s) P(ds) \quad (4.8)$$

gesetzt wurde. Damit konvergieren die endlich dimensionalen Verteilungen von B_n unter $P_n^{\otimes n}$ gegen die endlich dimensionalen Verteilungen von $B + A$. Aus der Darstellung (4.5) ergibt sich nach dem Lemma 5.1 aus dem Anhang I, dass die Folge $P_n^{\otimes n}$ zur Folge $P^{\otimes n}$ benachbart ist.

Hieraus folgt

$$\mathcal{L}(B_n | P_n^{\otimes n}) \triangleleft \mathcal{L}(B_n | P^{\otimes n}).$$

Aus dem Lemma 5.7 im Anhang II erhalten wir folgendes Resultat, das bereits in etwas anderer Formulierung auf der Seite 167 in Shorack und Wellner (1986) zu finden ist.

Satz 4.4 *Ist die Verteilungsfunktion F stetig, P die zugehörige Verteilung und sind die Voraussetzungen (4.4) und (4.6) erfüllt, so gilt*

$$\mathcal{L}(B_n | P_n^{\otimes n}) \Longrightarrow \mathcal{L}(B(F) + A). \quad (4.9)$$

Aus dem stetigen Abbildungssatz und der Tatsache, dass $\kappa(x) = \sup_{-\infty \leq t \leq \infty} |x(t)|$ eine stetige und messbare Funktion auf $\mathbb{D}[-\infty, \infty]$ ist, erhalten wir

$\mathcal{L}(\kappa(B_n) | P_n^{\otimes n}) \Longrightarrow \mathcal{L}(\kappa(B(F) + A))$. Eine analoge Aussage für die Cramer-von-Mises-Statistik (4.2) ergibt sich aus Lemma 5.8 im Anhang II.

Proposition 4.5 *Unter den Voraussetzungen von Satz 4.4 gelten*

$$\begin{aligned} \mathcal{L}\left(\sup_{-\infty \leq t \leq \infty} |B_n(t)| | P_n^{\otimes n}\right) &\Longrightarrow \mathcal{L}\left(\sup_{-\infty \leq t \leq \infty} |B(F(t)) + A(t)|\right), \\ \mathcal{L}\left(\sum_{i=1}^n \left(\frac{i}{n} - F(X_{n,i})\right)^2 | P_n^{\otimes n}\right) &\Longrightarrow \mathcal{L}\left(\int (B(F(t)) + A(t))^2 P(dt)\right). \end{aligned}$$

4 Anpassungstests

Ist P die Verteilung auf der reellen Achse mit der Verteilungsfunktion $F(t) = P((-\infty, t])$ und ist F stetig, so gilt

$$\int_{-\infty}^{\infty} g(F(t))P(dt) = \int_0^1 g(s)ds, \quad (4.10)$$

für jedes $g \in L_1[0, 1]$, siehe Seite 20-21 in Witting und Müller-Funk (1995). Weiter gilt

$$P(\{x : F^{-1}(F(x)) = x\}) = 1.$$

Damit folgt für jedes $h \in L_1(P)$

$$\int_{-\infty}^{\infty} h(t)P(dt) = \int_0^1 h(F^{-1}(s))ds. \quad (4.11)$$

Das ergibt

$$\int (B(F(t)) + A(t))^2 P(dt) = \int (B(s) + \mathcal{A}(s))^2 ds$$

mit $\mathcal{A}(s) = A(F^{-1}(s))$. Aus der Kac-Siebert-Zerlegung in (5.18) im Anhang III ergibt sich mit $\varphi_k(s) = \sqrt{2} \sin(k\pi s)$

$$B(s) + \mathcal{A}(s) = \sum_{k=1}^{\infty} (Z_k + A_k k\pi) \varphi_k(s) \quad (4.12)$$

$$\text{mit } A_k = \int_0^1 \mathcal{A}(s) \varphi_k(s) ds = \int A(t) \varphi_k(F(t)) P(dt).$$

Das ergibt

$$\mathcal{L}\left(\int (B(F(t)) + A(t))^2 P(dt)\right) = \mathcal{L}\left(\sum_{k=1}^{\infty} \frac{1}{k^2 \pi^2} (Z_k + A_k k\pi)^2\right).$$

Wir betrachten jetzt für die Klasse von stochastischen Prozessen

$$\begin{aligned} B(F(t)) + A(t), \\ A(t) = \int I_{(-\infty, t]}(s) a(s) P(ds) \end{aligned}$$

das Testproblem $H_0 : a = 0$ gegen $H_A : a \neq 0$, $a \in L_2^0(P)$ und verwenden dazu die Teststatistik

$$T = \int B^2(t) P(dt).$$

Wir lehnen die Nullhypothese für große Werte von T ab. Zu fest gewähltem $0 < \alpha < 1$ sei $c_{1-\alpha}$ das α -Quantil der Verteilung von $\sum_{k=1}^{\infty} \frac{1}{k^2 \pi^2} Z_k^2$. Dann ist die Gütefunktion dieses α -Tests gegeben durch

$$g(\alpha, a) = P\left(\sum_{k=1}^{\infty} \frac{1}{k^2 \pi^2} (Z_k + A_k k\pi)^2 \geq c_{1-\alpha}\right). \quad (4.13)$$

Wir formen zunächst A_k etwas um. Es gilt

$$\begin{aligned}
 A_k &= \int \left(\int I_{(-\infty, s]}(\tau) a(\tau) P(d\tau) \right) \sqrt{2} \sin(k\pi F(s)) P(ds) \\
 &= \int \left(\int I_{[F(\tau), 1]}(F(s)) \sqrt{2} \sin(k\pi F(s)) P(ds) \right) a(\tau) P(d\tau) \\
 &= \int \left(\int I_{[F(\tau), 1]}(u) \sqrt{2} \sin(k\pi u) du \right) a(\tau) P(d\tau) \\
 &= \int \left(-\frac{\sqrt{2}}{k\pi} \cos(k\pi u) \Big|_{F(\tau)}^1 \right) a(\tau) P(d\tau) \\
 &= \int \left(\frac{\sqrt{2}}{k\pi} \cos(k\pi F(\tau)) \right) a(\tau) P(d\tau),
 \end{aligned}$$

wobei die letzte Beziehung aus $\int a(\tau) P(d\tau) = 0$ folgt. Somit haben wir

$$\begin{aligned}
 A_k &= \frac{1}{k\pi} a_k, \\
 a_k &= \int a(\tau) \left(\sqrt{2} \cos(k\pi F(\tau)) \right) P(d\tau)
 \end{aligned} \tag{4.14}$$

erhalten.

Gehen wir von der Funktion a zur Funktion ρa über, so ist die neue Gütefunktion gegeben durch

$$g(\alpha, \rho a) = P \left(\sum_{k=1}^{\infty} \frac{1}{k^2 \pi^2} (Z_k + \rho a_k)^2 \geq c_{1-\alpha} \right).$$

Weil für eine zentrierte normalverteilte Zufallsvariable Z die Zufallsvariablen $(Z + A)^2$ stochastisch wachsend in A^2 sind und die Zufallsvariablen Z_k unabhängig sind, ergibt sich, dass $\sum_{k=1}^{\infty} \frac{1}{k^2 \pi^2} (Z_k + \rho a_k)^2$ stochastisch wachsend in ρ^2 ist. Bei einem Vergleich verschiedener lokaler Alternativen, müssen wir deshalb die entsprechenden Funktionen aus $L_2^0(P)$ normieren. Der folgende Satz geht in anders formulierter Version auf Neuhaus (1976) zurück und beantwortet die Frage, für welche Alternativen die größte Güte erreicht wird.

Satz 4.6 Für jedes $a \in L_2^0(P)$ gilt für die Gütefunktion des Cramer-von-Mises-Tests

$$\begin{aligned}
 g(\alpha, a) &= P \left(\sum_{k=1}^{\infty} \frac{1}{k^2 \pi^2} (Z_k + a_k)^2 \geq c_{1-\alpha} \right), \\
 a_k &= \int a(\tau) \left(\sqrt{2} \cos(k\pi F(\tau)) \right) P(d\tau).
 \end{aligned}$$

Sei $a^*(t) = \sqrt{2} \cos(\pi F(t))$. Dann gilt $a^* \in L_2^0(P)$ und für alle $a \in L_2^0(P)$, $0 < \alpha < 1$

$$g(\alpha, a) \leq g(\alpha, \rho a^*),$$

wobei $\rho^2 = \int a^2(t) P(dt)$ gesetzt wurde.

Wir wollen jetzt die oben bewiesenen Resultate zur Konstruktion von Anpassungstest anwenden. Gegeben sei eine Verteilung P , die auf den Borelmengen der reellen Achse definiert ist. $F(x) = P((-\infty, x])$ sei die zugehörige Verteilungsfunktion, die als stetig angenommen

4 Anpassungstests

wird. Für eine Stichprobe vom Umfang n , d. h. für n u.i.v. Zufallsvariablen X_1, \dots, X_n wird folgendes Testproblem betrachtet

$$H_0 : \mathcal{L}(X_i) = P \quad \text{gegen} \quad H_A : \mathcal{L}(X_i) \neq P. \quad (4.15)$$

Zur Beurteilung der Güte von Anpassungstests untersucht man, wie sich diese Güte verhält, wenn die Alternative aus einer einparametrischen Kurve im Raum der Verteilungen kommt, die in einer bestimmten Richtung von der Verteilung P abweicht. In diesem Zusammenhang spielt der aus der Literatur bekannte Begriff der L_2 -Ableitung oder der Tangente einer einparametrischen Verteilungsfamilie eine entscheidende Rolle, siehe Anhang I. Die einparametrische Familie $(Q_\eta)_{-\varepsilon < \eta < \varepsilon}$ sei L_2 -differenzierbar mit Ableitung $\dot{L}_0 \in L_2(Q_0)$. Wegen (5.7) gilt

$$\int \dot{L}_0 dQ_0 = 0.$$

Ist F stetig, so folgen mit $b(s) = \dot{L}_0(F^{-1}(s))$ aus (4.10)

$$\begin{aligned} \int_0^1 b(s) ds &= \int \dot{L}_0 dQ_0 = 0 \quad \text{und} \\ \int_0^1 b^2(s) ds &= \int \dot{L}_0^2 dQ_0 < \infty. \end{aligned}$$

Aus Lemma 5.2 im Anhang II erhalten wir, dass bei stetiger Verteilungsfunktion F jede Funktion $b(F)$ mit $b \in L_2^0[0, 1]$ die Tangente \dot{L}_0 einer L_2 -differenzierbaren Kurve ist.

Sei $(Q_\eta)_{\eta \in \Delta}$ eine L_2 -differenzierbare Kurve mit $Q_0 = P$ mit Ableitung $\dot{L}_0 = b(F)$. Wir setzen $P_n = Q_{1/\sqrt{n}}$. Dann ist die Bedingung (4.4) mit $a = b(F_0)$ erfüllt.

Die folgende Aussage ist eine direkte Konsequenz von Proposition 4.5 und Satz 4.6.

Proposition 4.7 *Gegeben sei das Testproblem (4.15). Sei $(Q_\eta)_{\eta \in \Delta}$ eine L_2 -differenzierbare Kurve mit $Q_0 = P$ und der Tangente $\dot{L}_0 = b(F)$, die die Länge ρ besitzt, d. h. es gilt*

$$\int \dot{L}_0^2 dQ = \int_0^1 b^2(s) ds = \rho^2.$$

Die Verteilungsfunktion $F(t) = P((-\infty, t])$ sei stetig. Für jedes feste $0 < \alpha < 1$ wird die asymptotische Güte $g(\alpha, b(F))$ des Cramer-von-Mises-Tests in (4.13) maximal, wenn $b(t) = \sqrt{2} \cos(\pi t)$ gilt, d. h. das Testproblem (4.15) in die spezielle Folge von Testproblemen

$$H_0 : \mathcal{L}(X_{n,i}) = P \quad \text{gegen} \quad H_A : \mathcal{L}(X_{n,i}) = Q_n$$

übergeht, wobei

$$\frac{dQ_n}{dP} = 1 + \frac{\rho}{\sqrt{n}} \sqrt{2} \cos(\pi F)$$

gilt.

Hat P die Lebesguedichte f , so hat Q_n die Lebesguedichte

$$f_n = \left(1 + \frac{\rho}{\sqrt{n}} \sqrt{2} \cos(\pi F)\right) f.$$

4.1.3 Konvergenz des empirischen Prozesses mit geschätzten Parametern

Wir betrachten jetzt eine parametrisierte Familie $(P_\theta)_{\theta \in \Theta}$ von Verteilungen und wollen testen, ob die Verteilung der u.i.v. X_1, \dots, X_n aus dem Modell $(P_\theta)_{\theta \in \Theta}$ stammt. Zu diesem Zweck setzen wir $F_\theta(t) = P_\theta((-\infty, t])$ und die empirische Verteilungsfunktion \hat{F}_n mit $F_{\hat{\theta}_n}$, wobei $\hat{\theta}_n$ eine konsistente Schätzfolge ist. Das führt uns zum geschätzten empirischen Prozess, der durch

$$\hat{B}_n = \sqrt{n}(\hat{F}_n - F_{\hat{\theta}_n})$$

definiert ist. Die asymptotische Verteilung \hat{B}_n wurde zuerst von Durbin (1973a) und Durbin (1973b) untersucht und später von verschiedenen Autoren durch Abschwächung der Voraussetzungen verallgemeinert und modifiziert. Die verschiedenen Resultate in der Literatur unterscheiden sich hinsichtlich der Regularitätsbedingungen, die notwendig sind, um eine geeignete Taylorentwicklung zu machen, siehe Shorack und Wellner (1986), Kapitel 5.5. Unser Zugang folgt van der Vaart (1998) und Genz und Haeusler (2006). Wir nehmen an, dass Θ eine offene Teilmenge von \mathbb{R}_d ist und die Folge der Schätzer $\hat{\theta}_n$ konsistent ist, d. h.

$$\hat{\theta}_n(X_1, \dots, X_n) \xrightarrow{P_\theta^{\otimes n}} \theta. \quad (4.16)$$

Weiterhin setzen wir voraus, dass $\hat{\theta}_n$ eine geeignete asymptotische Darstellung besitzt. Es möge eine messbare Funktion $h_\theta : \mathbb{R} \rightarrow \mathbb{R}_d$ derart existieren, dass gilt

$$\sqrt{n}(\hat{\theta}_n(X_1, \dots, X_n) - \theta) = \frac{1}{\sqrt{n}} \sum_{i=1}^n h_\theta(X_i) + o_{P_\theta^{\otimes n}}(1), \quad (4.17)$$

$$E_\theta \|h_\theta(X_1)\|^2 < \infty, \quad E_\theta h_\theta(X_1) = 0, \quad (4.18)$$

$$J(\theta) = E_\theta h_\theta(X_1) h_\theta^T(X_1). \quad (4.19)$$

Hierbei bezeichnet $\|x\|$ die Euklidische Norm des Spaltenvektors x und \top ist das Symbol für die Transposition. Wir setzen voraus, dass für jedes $\theta \in \Theta$ eine Funktion \dot{F}_θ existiert, die als Funktion von (θ, t) stetig ist, so dass gelten

$$\sup_{t \in [-\infty, \infty]} |F_{\theta+h}(t) - F_\theta(t) - \dot{F}_\theta^T(t)h| = o(\|h\|) \quad \text{falls } h \rightarrow 0, \quad (4.20)$$

$$\dot{F}_\theta \in \mathbb{C}[-\infty, \infty] \text{ für alle } \theta \text{ und } (\theta, t) \mapsto \dot{F}_\theta(t) \text{ ist stetig.}$$

Falls $\theta \mapsto F_\theta(t)$ differenzierbar ist, dann ist der Gradient

$$\dot{F}_\theta(t) = \left(\frac{\partial}{\partial \theta_1} F_\theta(t), \dots, \frac{\partial}{\partial \theta_d} F_\theta(t) \right)^T \quad (4.21)$$

ein typischer Kandidat für $\dot{F}_\theta(t)$. Der entscheidende Schritt zum Nachweis eines Grenzwertsatzes für $\hat{B}_n(t)$ ist die folgende Linearisierung, die z. B. in van der Vaart (1998) und Genz und Haeusler (2006) zu finden ist. Weil diese Linearisierung grundlegend ist, sollen alle Details des Beweises hier nochmals angegeben werden.

Lemma 4.8 *Unter den Voraussetzungen (4.17), (4.18) und (4.20) gilt*

$$\sup_t \left| \hat{B}_n(t) - n^{-1/2} \sum_{i=1}^n (I_{(-\infty, t)}(X_i) - F_\theta(t) - \dot{F}_\theta^T(t)h_\theta(X_i)) \right| = o_{P_\theta^{\otimes n}}(1). \quad (4.22)$$

Beweis. Sei

$$R(\varepsilon) = \sup_{\|h\| \leq \varepsilon} \sup_{t \in [-\infty, \infty]} |F_{\theta+h}(t) - F_{\theta}(t) - \dot{F}_{\theta}^T(t)h|.$$

Dann gilt nach Voraussetzung (4.20) die Beziehung

$$\lim_{\varepsilon \rightarrow 0} R(\varepsilon)/\varepsilon = 0. \quad (4.23)$$

Sei $A_n(\varepsilon) = \left\{ \left\| \hat{\theta}_n - \theta \right\| > c/\sqrt{n} \right\}$. Dann folgt

$$\begin{aligned} & \left| \hat{B}_n(t) - n^{-1/2} \sum_{i=1}^n (l_{(-\infty, t)}(X_i) - F_{\theta}(t) - \dot{F}_{\theta}^T(t)h_{\theta}(X_i)) \right| \\ & \leq \left| n^{1/2}(F_{\hat{\theta}_n}(t) - F_{\theta}(t)) - \dot{F}_{\theta}^T(t)n^{-1/2} \sum_{i=1}^n h_{\theta}(X_i) \right| \\ & \leq n^{1/2} \left| (F_{\hat{\theta}_n}(t) - F_{\theta}(t)) - \dot{F}_{\theta}^T(t)(\hat{\theta}_n - \theta) \right| + o_{P_{\theta}^{\otimes n}}(1), \end{aligned}$$

wobei wir für die letzte Ungleichung

$$C := \sup_t \left\| \dot{F}_{\theta}(t) \right\| < \infty$$

verwendet haben. Weiterhin gilt

$$\begin{aligned} & \sup_t n^{1/2} \left| (F_{\hat{\theta}_n}(t) - F_{\theta}(t)) - \dot{F}_{\theta}^T(t)(\hat{\theta}_n - \theta) \right| \leq \\ & \leq c \frac{R(c/\sqrt{n})}{c/\sqrt{n}} + n^{1/2} \sup_t \left| (F_{\hat{\theta}_n}(t) - F_{\theta}(t)) - \dot{F}_{\theta}^T(t)(\hat{\theta}_n - \theta) \right| I_{A_n(c/\sqrt{n})}. \end{aligned} \quad (4.24)$$

Weil

$$\begin{aligned} E_{\theta} \left\| \frac{1}{\sqrt{n}} \sum_{i=1}^n h_{\theta}(X_i) \right\|^2 &= \frac{1}{n} \sum_{i=1}^n E_{\theta} h_{\theta}^T(X_i) h_{\theta}(X_i) \\ &= E_{\theta} \|h_{\theta}(X_1)\|^2 < \infty \end{aligned}$$

gilt, ist die Folge $\frac{1}{\sqrt{n}} \sum_{i=1}^n h_{\theta}(X_i)$ stochastisch beschränkt. Dann ist auch

$$\sqrt{n}(\hat{\theta}_n - \theta) = \frac{1}{\sqrt{n}} \sum_{i=1}^n h_{\theta}(X_i) + o_{P_{\theta}^{\otimes n}}(1)$$

stochastisch beschränkt, d. h. es gilt

$$\lim_{c \rightarrow \infty} \limsup_{n \rightarrow \infty} P_{\theta}^{\otimes n} \left(\left\| \sqrt{n}(\hat{\theta}_n - \theta) \right\| \geq c \right) = 0.$$

Dann kann man eine hinreichend langsam gegen Unendlich strebende Folge $c_n \rightarrow \infty$ derart finden, dass gilt

$$\limsup_{n \rightarrow \infty} P_{\theta}^{\otimes n} \left(\left\| \sqrt{n}(\hat{\theta}_n - \theta) \right\| \geq c_n \right) = 0.$$

Durch Übergang zu einer noch langsamer gegen Unendlich strebenden Folge können wir wegen (4.23) erreichen, dass gilt

$$\lim_{n \rightarrow \infty} c_n \frac{R(c_n/\sqrt{n})}{c_n/\sqrt{n}} = 0.$$

Ersetzen wir in der rechten Seite von (4.24) c durch die Folge c_n , so strebt die rechte Seite stochastisch gegen Null und die Behauptung ist bewiesen. ■

Satz 4.9 Unter den Voraussetzungen von Satz 4.4 und Lemma 5.2 gilt

$$\mathcal{L}(\widehat{B}_n | P_n^{\otimes n}) \Rightarrow \mathcal{L}(Z + A - A_\theta)$$

im Sinne der schwachen Konvergenz von Verteilungen auf $\mathbb{D}[-\infty, \infty]$. Hierbei ist

$$A_\theta(t) = \dot{F}_\theta^T(t) \int a(x) h_\theta(x) P_\theta(dx)$$

und Z ist ein stetiger Gaußprozess mit Erwartungswert Null und der Kovarianzfunktion

$$\begin{aligned} K_\theta(s, t) = & F_\theta(s \wedge t) - F_\theta(s)F_\theta(t) + \dot{F}_\theta^T(s)J(\theta)\dot{F}_\theta(t) \\ & - \dot{F}_\theta^T(s)Eh_\theta(X_1)I_{(-\infty, t]}(X_1) - \dot{F}_\theta^T(t)Eh_\theta(X_1)I_{(-\infty, s]}(X_1) \end{aligned} \quad (4.25)$$

für alle $s, t \in [-\infty, \infty]$.

Folgerung 4.10 Unter den Voraussetzungen von Satz 4.9 gilt

$$\begin{aligned} \mathcal{L}\left(\sup_t |\widehat{B}_n(t)| | P_n^{\otimes n}\right) &\Rightarrow \mathcal{L}\left(\sup_t |Z(t) + A(t) - A_\theta(t)|\right), \\ \mathcal{L}\left(\int (\widehat{B}_n(t))^2 d\widehat{F}_n(t) | P_n^{\otimes n}\right) &\Rightarrow \mathcal{L}\left(\int (Z(t) + A(t) - A_\theta(t))^2 P_\theta(dt)\right). \end{aligned}$$

Beweis. Wir setzen

$$\begin{aligned} Z_n(t) &= Z_{1,n}(t) + Z_{2,n}(t) \\ &= n^{-1/2} \sum_{i=1}^n (I_{(-\infty, t)}(X_{n,i}) - F_\theta(t)) - n^{-1/2} \dot{F}_\theta^T(t) \sum_{i=1}^n h_\theta(X_{n,i}). \end{aligned}$$

Seien t_1, \dots, t_m und $\alpha_1, \dots, \alpha_m$ fest gewählt. Wir setzen

$$b(x) = \sum_{j=1}^m \alpha_j (I_{(-\infty, t_j)}(x) - F_\theta(t_j) - \dot{F}_\theta^T(t_j) h_\theta(x))$$

und wenden Proposition 5.5 im Anhang I an. Das ergibt

$$\mathcal{L}\left(\sum_{j=1}^m \alpha_j Z_n(t_j) | P_n^{\otimes n}\right) \Rightarrow \mathcal{N}\left(\int a(x)b(x)P_\theta(dx), \int b^2(x)P_\theta(dx)\right).$$

Wegen $\int a(x)P_\theta(dx) = 0$ und $A(t) = \int a(x)I_{(-\infty, t)}(x)P_\theta(dx)$ nach (4.8) folgen

$$\begin{aligned} \int a(x)b(x)P_\theta(dx) &= \sum_{j=1}^m \alpha_j (A(t_j) - \dot{F}_\theta^T(t_j) \int h_\theta(x)a(x)P_\theta(dx)) \\ &= \sum_{j=1}^m \alpha_j (A(t_j) - A_\theta(t_j)), \\ \int b^2(x)P_\theta(dx) &= \sum_{j=1}^m \alpha_j \alpha_j K_\theta(t_i, t_j). \end{aligned}$$

Damit haben wir die Konvergenz der endlich dimensionalen Verteilungen von Z_n gegen die entsprechenden endlich dimensionalen Verteilungen von $Z + A - A_\theta$ gezeigt. Aus (4.22) folgt, dass die endlich dimensionalen Verteilungen von B_n gegen die entsprechenden endlich dimensionalen Verteilungen von Z konvergieren. Wegen $\mathcal{L}(Z_{1,n}) \Rightarrow \mathcal{L}(B(F_\theta))$ nach Satz 4.1 und Lemma 5.6 erfüllt die Folge der zufälligen Element $Z_{1,n}$ die Bedingung (5.11). Weil die Folge $n^{-1/2} \sum_{i=1}^n h_\theta(X_{n,i})$ stochastisch beschränkt ist, folgt die Bedingung (5.11) für $Z_{2,n}$. Dann erfüllt $Z_{1,n} + Z_{2,n}$ auch diese Bedingung. Aus (4.22) ergibt sich dann, dass \widehat{B}_n auch die Bedingung (5.11) erfüllt. Das ergibt zusammen mit der bereits bewiesenen Konvergenz der endlich dimensionalen Verteilungen die Behauptung. Die erste Behauptung in der Folgerung ergibt sich wieder aus der Tatsache, dass κ in (4.1) eine stetige, beschränkte und $\mathfrak{B}_{\mathbb{D}}$ -messbare Funktion ist. Die zweite Behauptung folgt aus Lemma 5.8 im Anhang II. ■

4.2 Der randomisierte empirische Prozesses

Weil die Kovarianzfunktion des Prozesses Z im Satz 4.9 vom unbekannten Parameter abhängt, sind der Kolmogoroff-Smirnov-Test und der Cramer-von-Mises-Test jetzt nicht mehr verteilungsfrei. Die entsprechenden Quantile hängen vom gesamten parametrischen Modell und innerhalb des Modells auch noch vom unbekannten Parameter θ ab. Wir werden mit Hilfe einer neuen Randomisierungstechnik diese Abhängigkeit von θ beseitigen. Dazu betrachten wir den Spezialfall, wenn $\hat{\theta}_n$ der Maximum-Likelihood-Schätzer ist. Wir nehmen an, dass $(P_\theta)_{\theta \in \Theta}$ durch ein σ -endliches Maß μ ohne diskrete Anteile dominiert wird und setzen $f_\theta = dP_\theta/d\mu$, $\theta \in \Theta$. Weil μ atomlos ist, ist die Verteilungsfunktion

$$F_\theta(t) = \int_{-\infty}^t f_\theta(s) \mu(ds) \quad (4.26)$$

stetig in t . Wir setzen voraus, dass folgende Regularitätsbedingungen erfüllt sind

$$\begin{aligned} f_\theta(x) &= \frac{dP_\theta}{d\mu}(x) > 0, \quad \mu - a.s \text{ für alle } \theta \in \Theta \\ \theta &\mapsto f_\theta(x) \text{ ist stetig differenzierbar für alle } x \\ \int \|\dot{f}_\theta(x)\|^2 \frac{1}{f_\theta(x)} \mu(dx) &< \infty \\ \theta &\rightarrow I(\theta) = \int \dot{f}_\theta(x) \dot{f}_\theta^T(x) \frac{1}{f_\theta(x)} \mu(dx) \text{ ist stetig,} \\ \det(I(\theta)) &\neq 0 \text{ für jedes } \theta \in \Theta \\ \int_{-\infty}^t \dot{f}_\theta(x) \mu(dx) &= \dot{F}_\theta(t), \quad -\infty < t < \infty. \end{aligned} \quad (4.27)$$

wobei $\dot{f}_\theta := (\frac{\partial}{\partial \theta_1} f_\theta, \dots, \frac{\partial}{\partial \theta_d} f_\theta)^T$ ist. Die letzte Bedingung in (4.27) bedeutet, dass man die Ableitung nach θ mit dem Integral nach s in (4.26) vertauschen kann. $I(\theta)$ ist die Fisher-Informationsmatrix. Wir bemerken zunächst, dass die obigen Bedingungen die L_2 -Differenzierbarkeit des Modells nach sich ziehen, wobei wir für den Beweis auf Liese und Miescke (2008) verweisen.

Lemma 4.11 *Sind die Bedingungen in (4.27) erfüllt, so ist das Modell $(P_\theta)_{\theta \in \Theta}$ in jedem inneren Punkt von Θ im Sinne von Definition 5.3 L_2 -differenzierbar mit der Ableitung*

$$\begin{aligned} \dot{L}_\theta &= \frac{1}{f_\theta(x)} \dot{f}_\theta \\ &= \frac{1}{f_\theta} \left(\frac{\partial}{\partial \theta_1} f_\theta, \dots, \frac{\partial}{\partial \theta_d} f_\theta \right)^T. \end{aligned} \quad (4.28)$$

Wir betrachten jetzt die Kovarianzfunktion im Satz 4.9. Es gilt

$$\begin{aligned} \dot{F}_\theta^T(s) E h_\theta(X_1) I(X_1 \leq t) &= \dot{F}_\theta(s)^T \int_{-\infty}^t I^{-1}(\theta) \dot{L}_\theta(x) P_\theta(dx) \\ &= \dot{F}_\theta(s)^T \int_{-\infty}^t I^{-1}(\theta) \frac{1}{f_\theta(x)} \dot{f}_\theta(x) f_\theta(x) \mu(dx) = \dot{F}_\theta(s)^T I^{-1}(\theta) \int_{-\infty}^t \dot{f}_\theta(x) \mu(dx) \\ &= \dot{F}_\theta(s)^T I^{-1}(\theta) \dot{F}_\theta(t), \end{aligned}$$

wobei die letzte Gleichung aus der letzten Bedingung in (4.27) folgt. Damit hat die Kovarianzfunktion (4.25) die Gestalt

$$\text{cov}(Z(s), Z(t)) = F_\theta(s \wedge t) - F_\theta(s) F_\theta(t) - \dot{F}_\theta^T(t) I^{-1}(\theta) \dot{F}_\theta(s). \quad (4.29)$$

Aus dem Satz 4.9 erhalten wir sofort folgendes Ergebnis.

Satz 4.12 Es seien die Voraussetzungen von Satz 4.4 und Lemma 5.2 mit $h_\theta = I^{-1}(\theta)\dot{L}_\theta$ erfüllt. $\hat{\theta}_n$ sei der Maximum-Likelihood-Schätzer und es seien die Bedingungen in (4.27) erfüllt. Dann gilt

$$\mathcal{L}(\widehat{B}_n | P_n^{\otimes n}) \Rightarrow \mathcal{L}(Z + A - A_\theta)$$

im Sinne der schwachen Konvergenz von Verteilungen auf $\mathbb{D}[-\infty, \infty]$. Hierbei ist

$$A_\theta(t) = \dot{F}_\theta^T(t) \int a(x) I^{-1}(\theta) \dot{L}_\theta(x) P_\theta(dx)$$

und Z ist ein stetiger Gaußprozess mit Erwartungswert Null und der Kovarianzfunktion (4.29).

Die Kovarianzmatrix $\text{cov}(Z(t_i), Z(t_j))$, $i, j = 1, \dots, n$ ist positiv semidefinit. Sie ist zugleich die Differenz der beiden Matrizen $F_\theta(t_i \wedge t_j) - F_\theta(t_i)F_\theta(t_j)$ und $\dot{F}_\theta^T(t_i)I^{-1}(\theta)\dot{F}_\theta(t_j)$, die auch positiv semidefinit sind. Deshalb ist die Matrix $\text{cov}(Z(t_i), Z(t_j))$, $i, j = 1, \dots, n$ im Sinne der Loewner-Ordnung kleiner oder gleich der Matrix $F_\theta(t_i \wedge t_j) - F_\theta(t_i)F_\theta(t_j)$, $i, j = 1, \dots, n$. Aufgrund dieser Beziehung können wir den Zusatzterm $\dot{F}_\theta^T(t)I^{-1}(\theta)\dot{F}_\theta(s)$ auf der rechten Seite von (4.29) dadurch eliminieren, dass wir zu \widehat{B}_n einen geeigneten unabhängigen Prozess addieren. Nachfolgend bezeichnet $I^{-1/2}(\theta)$ die positiv semidefinite symmetrische Matrix mit $I^{-1/2}(\theta)I^{-1/2}(\theta) = I^{-1}(\theta)$. Für standardnormalverteilte Vektoren V_n , die von den Beobachtungen unabhängig sind und $\widehat{B}_n = \sqrt{n}(\widehat{F}_n - F_{\hat{\theta}_n})$ setzen wir

$$\widehat{BR}_n = \widehat{B}_n + \dot{F}_{\hat{\theta}_n}^T I^{-1/2}(\hat{\theta}_n) V_n \quad (4.30)$$

und bezeichnen \widehat{BR}_n als *randomisierten geschätzten empirischen Prozess*. Weiterhin seien

$$RC_n = \int \widehat{BR}_n^2(t) d\widehat{F}_n(t) \quad \text{und} \quad RK_n = \sup_t |\widehat{BR}_n(t)|.$$

RC_n bezeichnen wir als *randomisierte Cramer-von-Mises-Statistik* und RK_n als *randomisierte Kolmogoroff-Smirnov-Statistik*.

Satz 4.13 Es seien die Voraussetzungen von Satz 4.4 und Lemma 5.2 mit $h_\theta = I^{-1}(\theta)\dot{L}_\theta$ erfüllt und $\hat{\theta}_n$ sei der Maximum-Likelihood-Schätzer und es seien die Bedingungen in (4.27) erfüllt. Dann gilt

$$\mathcal{L}(\widehat{BR}_n | P_n^{\otimes n}) \Rightarrow \mathcal{L}(B(F_\theta) + A - A_\theta) \quad (4.31)$$

im Sinne der schwachen Konvergenz von Verteilungen auf $\mathbb{D}[-\infty, \infty]$. Hierbei sind P_n eine Folge mit der Tangente $a \in L_2^0(P_\theta)$ und

$$\begin{aligned} A(t) &= \int I_{(-\infty, t]}(s) a(s) P_\theta(ds), \\ A_\theta(t) &= \dot{F}_\theta^T(t) \int a(x) I^{-1}(\theta) \dot{L}_\theta(x) P_\theta(dx) \end{aligned} \quad (4.32)$$

und B die Brownsche Brücke auf $[0, 1]$ mit der Kovarianzfunktion $s \wedge t - st$.

Beweis. Wegen der Stetigkeit von $\theta \mapsto \dot{F}_\theta$, $\theta \mapsto I^{-1/2}(\theta)$ und der Konsistenz von $\hat{\theta}_n$ reicht es aus, die Summe $\widehat{B}_n + \dot{F}_{\hat{\theta}_n}^T I^{-1/2}(\hat{\theta}_n) V_n$ der unabhängigen Prozesse \widehat{B}_n und $\dot{F}_{\hat{\theta}_n}^T I^{-1/2}(\hat{\theta}_n) V_n$ zu betrachten. Sei V ein standardnormalverteilter Vektor. Dann gilt

$$\mathcal{L}(\dot{F}_{\hat{\theta}_n}^T I^{-1/2}(\hat{\theta}_n) V_n) = \mathcal{L}(\dot{F}_\theta^T I^{-1/2}(\theta) V)$$

und

$$\mathcal{L}(\widehat{B}_n + \dot{F}_\theta^T I^{-1/2}(\theta) V_n | P_n^{\otimes n}) \Rightarrow \mathcal{L}(Z + \dot{F}_\theta^T I^{-1/2}(\theta) V + A - A_\theta)$$

folgt aus Satz 4.12. Um zu zeigen, dass $Z + \dot{F}_\theta^T I^{-1/2}(\theta) V$ eine zeittransformierte Brownsche Brücke ist, berechnen wir die Kovarianzfunktion dieses Prozesses. Die Unabhängigkeit von Z und V sowie Formel (4.29) ergeben

$$\begin{aligned} & \text{cov}(Z(s) + \dot{F}_\theta^T(s) I^{-1/2}(\theta) V, Z(t) + \dot{F}_\theta^T(t) I^{-1/2}(\theta) V) \\ &= \text{cov}(Z(s), Z(t)) + \text{cov}(\dot{F}_\theta^T(s) I^{-1/2}(\theta) V, \dot{F}_\theta^T(t) I^{-1/2}(\theta) V) \\ &= \text{cov}(Z(s), Z(t)) + E(\dot{F}_\theta^T(s) I^{-1/2}(\theta) V)(\dot{F}_\theta^T(t) I^{-1/2}(\theta) V)^T \\ &= \text{cov}(Z(s), Z(t)) + \dot{F}_\theta^T(s) I^{-1}(\theta) \dot{F}_\theta(t) \\ &= F_\theta(s \wedge t) - F_\theta(s) F_\theta(t) - \dot{F}_\theta^T(t) I^{-1}(\theta) \dot{F}_\theta(s) + \dot{F}_\theta^T(s) I^{-1}(\theta) \dot{F}_\theta(t). \\ &= F_\theta(s) \wedge F_\theta(t) - F_\theta(s) F_\theta(t). \end{aligned} \tag{4.33}$$

■

Wir wollen jetzt die Struktur der Funktion $A - A_\theta$ näher analysieren und greifen hierzu auf den Begriff des Tangentialraumes und der Projektion auf den Tangentialraum zurück. Zunächst bemerken wir, dass nach Lemma 4.11 die Regularitätsbedingungen in (4.27) implizieren, dass das Modell $\mathcal{P} = (P_\theta)_{\theta \in \Theta}$, $\Theta \subseteq \mathbb{R}_d$ in den inneren Punkten von Θ L_2 -differenzierbar im Sinne von Definition 5.3 ist mit der Ableitung (4.28). Wie im Anhang I verstehen wir unter dem Tangentialraum $\mathcal{T}(\theta)$ des Modells im Punkt θ die Menge der Funktionen

$$\mathcal{T}(\theta) = \{h^\top \dot{L}_\theta, h \in \mathbb{R}_d\}.$$

$\mathcal{T}(\theta)$ ist ein abgeschlossener Unterraum von $L_2^0(P_\theta)$. Der Raum $L_2^0(P_\theta)$ ist nach Lemma 5.2 identisch mit dem Raum aller Tangenten von einparametrischen Kurven $(Q_\eta)_{-\varepsilon < \eta < \varepsilon}$ von Verteilungen durch P_θ , d. h. $Q_0 = P_\theta$. Es ist leicht zu sehen, dass der Tangentialraum gerade aus den Tangenten von differenzierbaren Kurven besteht, die im Modell liegen, d. h. $(Q_\eta)_{-\varepsilon < \eta < \varepsilon} \subseteq \mathcal{P}$. Wir erinnern, dass nach Lemma 5.4 die Projektion von $a \in L_2^0(P_\theta)$ auf das orthogonale Komplement des Tangentialraumes gegeben ist durch

$$\begin{aligned} \Pi_\theta^\perp a &= a - c^T(\theta) I^{-1}(\theta) \dot{L}_\theta, \\ c(\theta) &= \int a \dot{L}_\theta dP_\theta. \end{aligned} \tag{4.34}$$

Sind die Bedingungen in (4.27) erfüllt, dann folgt mit (4.28) aus (4.32)

$$\begin{aligned} A(t) - A_\theta(t) &= \int I_{(-\infty, t]}(s) a(s) P_\theta(ds) - \dot{F}_\theta^T(t) \int a(x) I^{-1}(\theta) \dot{L}_\theta(x) P_\theta(dx) \\ &= \int I_{(-\infty, t]}(s) a(s) f_\theta(s) \mu(ds) - \left(\int I_{(-\infty, t]}(s) \dot{L}_\theta(s) f_\theta(s) \mu(ds) \right)^T I^{-1}(\theta) \int a(x) \dot{L}_\theta(x) P_\theta(dx) \\ &= \int I_{(-\infty, t]}(s) (a(s) - c^T(\theta) I^{-1}(\theta) \dot{L}_\theta(s)) P_\theta(ds) \\ &= \int I_{(-\infty, t]}(s) (\Pi_\theta^\perp a)(s) P_\theta(ds). \end{aligned}$$

Bemerkung 4.14 Zwischen den Sätzen 4.13 und 4.4 gibt es einen bemerkenswerten Zusammenhang. Ersetzt man im Satz 4.4 die Funktion A durch die Differenz $A - A_\theta$ aus Satz 4.13 und den Prozess B_n durch den Prozess \widehat{BR}_n , so gehen die entsprechenden Aussagen

ineinander über. Wendet man jetzt zur Konstruktion einer Teststatistik ein Funktional, z. B. κ aus (4.1) auf \widehat{BR}_n , an, betrachtet man also die Statistiken $\kappa(\widehat{BR}_n)$, so erhält man einen Grenzwertsatz für $\kappa(\widehat{BR}_n)$, indem man \widehat{BR}_n durch B_n und A durch die Differenz $A - A_\theta$ ersetzt. Eine entsprechende Aussage erhält man für die Statistiken

$$RC_n = \int \widehat{BR}_n(t) d\widehat{F}_n(t),$$

wenn man Lemma 5.8 im Anhang verwendet. Damit erhalten wir ohne weiteren Beweis folgende Aussage

Folgerung 4.15 Unter den Voraussetzungen von Satz 4.13 gelten

$$\mathcal{L}(\widehat{BR}_n | P_n^{\otimes n}) \Rightarrow \mathcal{L}(B(F_\theta) + \int I_{(-\infty, \cdot]}(s)(\Pi_\theta^\perp a)(s)P_\theta(ds))$$

und

$$\begin{aligned} \mathcal{L}(RK_n | P_n^{\otimes n}) &\Rightarrow \mathcal{L}(\sup_t |B(F_\theta(t)) + \int I_{(-\infty, t]}(s)(\Pi_\theta^\perp a)(s)P_\theta(ds)|), \\ \mathcal{L}(RC_n | P_n^{\otimes n}) &\Rightarrow \mathcal{L}(\int (B(F_\theta(t)) + \int I_{(-\infty, t]}(s)(\Pi_\theta^\perp a)(s)P_\theta(ds))^2 P_\theta(dt)). \end{aligned}$$

Zu fest gewähltem $0 < \alpha < 1$ sei $c_{1-\alpha}$ das $1 - \alpha$ -Quantil der Verteilung

$$\mathcal{L}(\int_0^1 B^2(s)ds) = \mathcal{L}(\sum_{k=1}^n \frac{1}{(k\pi)^2} Z_k^2).$$

Wir führen den *randomisierten Cramer-von-Mises-Test* φ_{RC_n} ein durch

$$\varphi_{RC_n} = I_{(c_{1-\alpha}, \infty)}(RC_n). \quad (4.35)$$

Mit Hilfe des in der Bemerkung 4.14 geschilderten Zusammenhangs zwischen den Sätzen 4.13 und 4.4 ergibt sich aus Satz 4.6 folgende Aussage.

Satz 4.16 Es seien die Voraussetzungen von Satz 4.4 für die Folge P_n mit der Tangente $a \in L_2^0(P_\theta)$ erfüllt. Weiterhin mögen die Voraussetzungen von Lemma 5.2 mit $h_\theta = I^{-1}(\theta)\dot{L}_\theta$ gelten und $\hat{\theta}_n$ sei der Maximum-Likelihood-Schätzer. Es seien die Bedingungen in (4.27) erfüllt. Dann gilt für den randomisierten Cramer-von-Mises-Test φ_{RC_n}

$$\begin{aligned} \lim_{n \rightarrow \infty} \int \varphi_{RC_n} dP_n^{\otimes n} &= P \left(\sum_{k=1}^{\infty} \frac{1}{k^2 \pi^2} (Z_k + a_k)^2 \geq c_{1-\alpha} \right) \\ \text{mit } a_k &= \int (\Pi_\theta^\perp a)(\tau) \left(\sqrt{2} \cos(k\pi F_\theta(\tau)) \right) P_\theta(d\tau). \end{aligned} \quad (4.36)$$

Die obige Aussage im Satz 4.16 werden wir jetzt verwenden, um die Güte des Cramer-von-Mises-Tests für konkrete Verteilungsfamilien zu analysieren. Im Gegensatz zur Situation im Satz 4.6 lassen sich jetzt nicht so einfach diejenige Richtungen von Abweichungen vom Modell finden, die eine maximale Güte liefern. Wir bemerken zunächst, dass wir uns auf Funktionen a einschränken können, die senkrecht auf dem Tangentialraum des Modells stehen,

d. h. $\Pi_\theta a = 0$. Der Einfachheit halber schränken wir uns auf den Fall einparametrischer Modelle ein. Wir setzen

$$b_\theta(s) = \frac{\dot{f}_\theta}{f_\theta}(F_\theta^{-1}) \text{ und } b(s) = a(F_\theta^{-1}).$$

Dann lässt sich die Bedingung $\Pi_\theta a = 0$ schreiben als $a \perp \frac{\dot{f}_\theta}{f_\theta}$ in $L_2(P_\theta)$ oder im Hinblick auf (4.11)

$$0 = \int a \frac{\dot{f}_\theta}{f_\theta} dP_\theta = \int_0^1 b(s) b_\theta(s) ds. \quad (4.37)$$

Wir setzen

$$b_k = \sqrt{2} \int_0^1 b(s) \cos(k\pi s) ds \text{ und } b_{\theta,k} = \sqrt{2} \int_0^1 b_\theta(s) \cos(k\pi s) ds.$$

Dann lassen sich die a_k in (4.36) darstellen als

$$a_k = b_k - b_{\theta,k}.$$

Wegen der Orthogonalität der Funktionen $\sqrt{2} \cos(k\pi s)$ lässt sich die Bedingung (4.37) schreiben als

$$\sum_{k=1}^{\infty} b_k b_{\theta,k} = 0,$$

man beachte $b_0 = \int_0^1 b(s) ds = 0$. Wir wollen jetzt diejenige auf Länge 1 normierte Funktion $b \in L_2^0[0, 1]$ finden, die der Bedingung $b \perp b_\theta$ genügt und die Funktion

$$G(b_1, b_2, \dots) = P\left(\sum_{k=1}^{\infty} \frac{1}{k^2 \pi^2} (Z_k + b_k)^2 \geq c_{1-\alpha}\right) \quad (4.38)$$

maximiert. Entwickeln wir die Funktion b in eine Fourierreihe

$$b(s) = \sqrt{2} \sum_{k=1}^{\infty} b_k \cos(k\pi s),$$

so lässt sich das Problem auch so formulieren

$$\text{Gesucht sind reelle Zahlen } b_1, b_2, \dots \quad (4.39)$$

$$\text{mit } \sum_{k=1}^{\infty} b_k^2 = 1 \text{ und } \sum_{k=1}^{\infty} b_k b_{\theta,k} = 0,$$

die die Funktion $G(b_1, b_2, \dots)$ maximieren.

Dieses Maximierungsproblem lässt sich nur numerisch lösen. Durch Maximieren der Funktion G in (4.38) erhält man die Zahlen b_k , welche die Richtung maximaler Güte liefern. Für die Anwendung des randomisierten Anpassungstests betrachten wir nur die Cramer-von-Mises-Teststatistik. Für eine konkrete Untersuchung des Kolmogoroff-Smirnov-Tests inklusive Computersimulation siehe Liese und Liu (2011).

4.3 Der Cramer-von-Mises-Test für spezielle Modelle und spezielle Alternativen

4.3.1 Normalverteilung mit unbekanntem μ und bekanntem σ^2

Sei $N(\mu, \sigma^2)$ die Normalverteilung mit Erwartungswert μ und Varianz σ^2 . Φ und $\varphi(x)$ bezeichnen die Verteilungsfunktion und die Dichte von $N(0, 1)$. Wir nehmen an, dass σ^2

4 Anpassungstests

bekannt ist und können dann ohne Einschränkung $\sigma^2 = 1$ voraussetzen. Das Gaußsche Lokationsmodell ist dann $P_\mu = \mathbf{N}(\mu, 1)$, $\mu \in \mathbb{R}$ mit

$$F_\mu(t) = \Phi(t - \mu) \quad \text{und} \quad \dot{F}_\mu(t) = -\varphi(t - \mu).$$

Weiterhin gelten

$$\begin{aligned} \dot{L}_\mu(t) &= \frac{\partial}{\partial \mu} \ln \varphi(t - \mu) = t - \mu, \quad I(\mu) = 1, \\ F_\mu^{-1}(s) &= \Phi^{-1}(s) + \mu, \quad \dot{L}_\mu(F_\mu^{-1}(s)) = s. \end{aligned}$$

Deshalb besteht der Tangentialraum an der Stelle μ aus den Vielfachen von $\dot{L}_\mu(t) = t - \mu$. Weiter gilt mit $\mu = \theta$ für $c(\theta)$ in (4.34)

$$\begin{aligned} c(\mu) &= \int a \dot{L}_\mu dP_\mu = \int a(t)(t - \mu)\varphi(t - \mu) dt \\ &= \int a(s + \mu)s\varphi(s) ds. \end{aligned}$$

Damit folgt aus (4.34)

$$(\Pi_\mu^\perp a)(t) = a(t) - \left(\int a(s + \mu)s\varphi(s) ds \right) (t - \mu). \quad (4.40)$$

Die Bedingung (4.20) ist erfüllt. \bar{X}_n ist der Maximum-Likelihood-Schätzer für μ und die Voraussetzungen (4.17), (4.18) und (4.19) sind erfüllt mit $h_\theta(x) = x$ und $J(\mu) = I(\mu) = 1$. Weiterhin sind die Bedingungen in (4.27) erfüllt. Der geschätzte empirische Prozess ist gegeben durch

$$\widehat{B}_n(t) = \sqrt{n}(\widehat{F}_n(t) - \Phi(t - \bar{X}_n)).$$

Nach Satz 4.12 gilt $\mathcal{L}(\widehat{G}_n) \Rightarrow \mathcal{L}(Z_\mu)$, wobei Z_μ ein zentrierter Gaußprozess ist, dessen Kovarianzfunktion aus (4.29) folgt. Es gilt

$$\text{cov}(Z_\mu(s), Z_\mu(t)) = \Phi((s \wedge t) - \mu) - \Phi(s - \mu)\Phi(t - \mu) - \varphi(s - \mu)\varphi(t - \mu).$$

Die Cramer-von-Mises-Statistik ist gegeben durch

$$\text{RC}_n = n \int (\widehat{F}_n(t) - \Phi(t - \bar{X}_n))^2 d\widehat{F}_n(t) = \sum_{i=1}^n \left(\frac{i}{n} - \Phi(X_{n:i} - \bar{X}_n) \right)^2.$$

Der randomisierte geschätzte empirische Prozess $\widehat{BR}_n(t)$ in (4.30) ist gegeben durch

$$\widehat{BR}_n(t) = \sqrt{n}(\widehat{F}_n(t) - \Phi(t - \bar{X}_n)) - \varphi(t - \bar{X}_n) V_n,$$

wobei die standardnormalverteilten V_n von den Daten X_1, \dots, X_n unabhängig sind. Dann erhalten wir für die randomisierte Cramer-von-Mises-Statistik

$$\text{RC}_n = \int (\widehat{BR}_n(t))^2 d\widehat{F}_n(t) = \sum_{i=1}^n \left(\frac{i}{n} - \Phi(X_{n:i} - \bar{X}_n) - \frac{1}{\sqrt{n}} \varphi(X_{n:i} - \bar{X}_n) V_n \right)^2.$$

Jetzt sei P_n eine Folge von lokalen Alternativen mit der Tangente $a \in L_2^0(P_\mu)$. Aus Satz 4.16 ergibt sich dann

$$\begin{aligned} \lim_{n \rightarrow \infty} P_n^{\otimes n}(\varphi_{\text{RC}_n} = 1) &= P\left(\sum_{k=1}^n \frac{1}{(k\pi)^2} (Z_k + a_k)^2 > c_{1-\alpha}\right) \\ \text{mit } a_k &= \int (\Pi_\mu^\perp a)(t) \left(\sqrt{2} \cos(k\pi \Phi(t - \mu)) \right) \varphi(t - \mu) dt, \end{aligned} \quad (4.41)$$

4 Anpassungstests

wobei $\Pi_\mu^\perp a$ durch (4.40) gegeben ist. Zur Vereinfachung der Formeln nehmen wir jetzt ohne Einschränkung $\mu = 0$ an. Aus (4.40) folgen dann

$$(\Pi_0 a)(t) = \left(\int a(s) s \varphi(s) ds \right) \cdot t \quad \text{und} \quad \Pi_0^\perp a = a - \Pi_0 a. \quad (4.42)$$

Wir wollen die asymptotische Güte des Cramer-von-Mises-Tests für eine Folge von Alternativen untersuchen, die die folgenden vorgegebenen Tangenten haben

$$a_{l,h} = h \varphi_l(\Phi), \quad \text{wobei} \quad \varphi_l(s) = \sqrt{2} \cos(l\pi s). \quad (4.43)$$

Der zusätzliche Parameter kontrolliert den Abstand von der Nullhypothese. Weil φ_l beschränkt ist, lassen sich Folgen von Alternativen mit diesen vorgegebenen Tangenten einfach finden. Wir setzen

$$P_{l,n,h}(A) = \int_A \left(1 + \frac{h}{\sqrt{n}} \varphi_l(\Phi) \right) \varphi(t) dt, \quad -1/\sqrt{2} < \frac{h}{\sqrt{n}} < 1/\sqrt{2}. \quad (4.44)$$

Es ist einfach zu sehen, dass diese Folge von Verteilungen der Bedingung (4.4) genügt, wenn man $P_n = P_{l,n,h}$, $a_n = \frac{1}{\sqrt{n}} a_{l,h}$ und $a = a_{l,h}$ setzt. Die Verteilungen $P_{l,n,h}$ haben für $l = 1, 2, \dots$ die Dichten

$$p_{l,\eta}(t) = (1 + \eta \varphi_l(\Phi)) \varphi(t), \quad -1/\sqrt{2} < \eta = \frac{h}{\sqrt{n}} < 1/\sqrt{2}.$$

Falls $l = 2m$ gerade ist, dann folgt mit $\Phi(-t) = 1 - \Phi(t)$ die Symmetrie der Dichte, d. h. $p_{2m,\eta}(-t) = p_{2m,\eta}(t)$. Für ungerades $l = 2m + 1$ erhalten wir

$$\begin{aligned} p_{2m+1,\eta}(-t) &= (1 + \eta \sqrt{2} \cos((2m+1)\pi(1 - \Phi(t)))) \varphi(t) \\ &= (1 - \eta \sqrt{2} \cos((2m+1)\pi\Phi(t))) \varphi(t) = p_{2m+1,-\eta}(t). \end{aligned}$$

In den Abbildungen 4.1 und 4.2 sind die Dichten $p_{l,\eta}$ für spezielle l und η dargestellt.

Nullhypothese	Alternativhypothesen	
	$p_{1,0.5}; p_{1,-0.5}$	rot
$p_{l,0}$ schwarz $l = 1, 2, 3, 4$	$p_{2,0.5}; p_{2,-0.5}$	grün
	$p_{3,0.5}; p_{3,-0.5}$	blau
	$p_{4,0.5}; p_{4,-0.5}$	orange

Tabelle 4.1: Gewählte Dichten in Abhängigkeit von l und η

Wir untersuchen jetzt die asymptotische Güte des Cramer-von-Mises-Tests für die Folge von Alternativen $P_{l,n,h}$ in (4.44). Aus (4.42) ergibt sich mit $a = a_{l,h}$ für festes $l = 1, 2, \dots$ und $\mu = 0$

$$(\Pi_0^\perp a_{l,h})(t) = h \varphi_l(\Phi(t)) - ht \left(\int \varphi_l(\Phi(s)) s d\Phi(s) \right).$$

Sei

$$\gamma_l = \int s \varphi_l(\Phi(s)) d\Phi(s) = \int_0^1 \Phi^{-1}(t) \varphi_l(t) dt. \quad (4.45)$$

Dann sind bei festem l die a_k in (4.41) gegeben durch

$$a_{k,l} = h \int (\varphi_l(\Phi(t)) - \gamma_l t) \varphi_k(\Phi(t)) d\Phi(t).$$

4 Anpassungstests

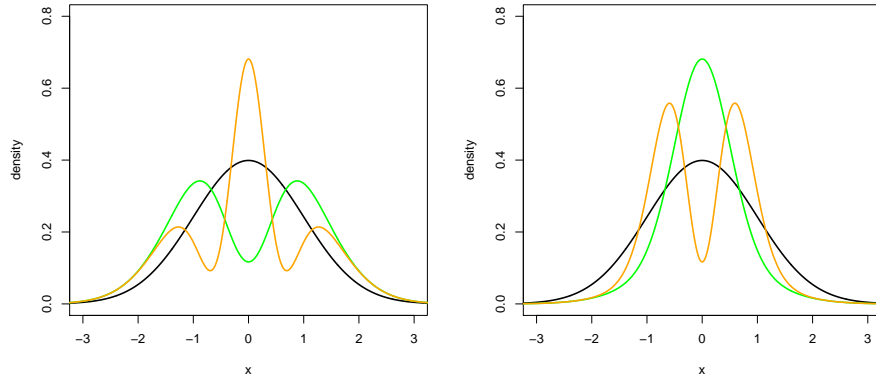


Abbildung 4.1: Verlauf der Dichte $p_{l,\eta}$. Linkes Bild: $\eta = 0.5$, $l = 2$ (grün); $\eta = 0.5$, $l = 4$ (orange). Rechtes Bild: $\eta = -0.5$, $l = 2$ (grün); $\eta = -0.5$, $l = 4$ (orange).

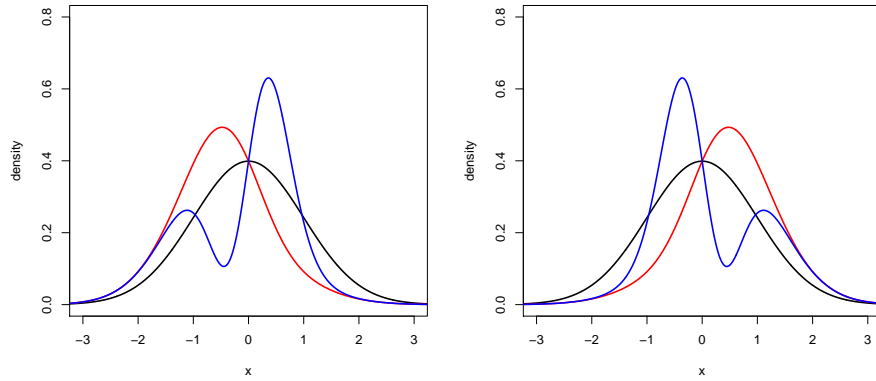


Abbildung 4.2: Verlauf der Dichte $p_{l,\eta}$. Linkes Bild: $\eta = 0.5$, $l = 1$ (rot); $\eta = 0.5$, $l = 3$ (blau). Rechtes Bild: $\eta = -0.5$, $l = 1$ (rot); $\eta = -0.5$, $l = 3$ (blau).

Weil die $\varphi_l(\Phi)$ ein Orthonormalsystem in $L_2(N(0, 1))$ bilden, gilt

$$a_{k,l} = h(\delta_{k,l} - \gamma_k \gamma_l), \quad (4.46)$$

wobei $\delta_{k,l}$ das Kroneckersymbol ist. Die Symmetrie der Standardnormalverteilung impliziert $\Phi^{-1}(0.5 + t) = -\Phi^{-1}(0.5 - t)$ für $0 \leq t \leq 0.5$ und

$$\int_0^1 \Phi^{-1}(t) g(t) dt = 0$$

für jede Funktion g , die symmetrisch bezüglich 0.5 ist. Damit erhalten wir aus (4.45)

$$\gamma_{2m} = \int_0^1 \Phi^{-1}(t) \varphi_{2m}(\Phi(t)) dt = 0. \quad (4.47)$$

Mit Hilfe von (4.46) und (4.47) erhalten wir die Statistik

$$S_l = \sum_{k=1}^{\infty} (k\pi)^{-2} (Z_k + h a_{k,l})^2,$$

4 Anpassungstests

die im Satz 4.16 auftritt, wenn man für $l = 1, \dots, 4$ in (4.43) die folgenden lokalen Alternativen betrachtet:

$$\begin{aligned}
 S_1 &= \frac{1}{\pi^2} (Z_1 + h(1 - \gamma_1^2))^2 + \sum_{m=1}^{\infty} \frac{1}{(2m\pi)^2} Z_{2m}^2 \\
 &\quad + \sum_{m=1}^{\infty} \frac{1}{((2m+1)\pi)^2} (Z_{2m+1} - h\gamma_1\gamma_{2m+1})^2, \\
 S_2 &= \frac{1}{\pi^2} Z_1^2 + \frac{1}{4\pi^2} (Z_2 + h)^2 + \sum_{k=3}^{\infty} \frac{1}{(k\pi)^2} Z_k^2, \\
 S_3 &= \frac{1}{\pi^2} (Z_1 - h\gamma_1\gamma_3)^2 + \frac{1}{4\pi^2} Z_2^2 + \frac{1}{9\pi^2} (Z_3 + h(1 - \gamma_3^2))^2 \\
 &\quad + \sum_{m=2}^{\infty} \frac{1}{(2m\pi)^2} Z_{2m}^2 + \sum_{m=2}^{\infty} \frac{1}{((2m+1)\pi)^2} (Z_{2m+1} - h\gamma_3\gamma_{2m+1})^2, \\
 S_4 &= \frac{1}{\pi^2} Z_1^2 + \frac{1}{4\pi^2} Z_2^2 + \frac{1}{9\pi^2} Z_3^2 + \frac{1}{16\pi^2} (Z_4 + h)^2 + \sum_{k=5}^{\infty} \frac{1}{(k\pi)^2} Z_k^2.
 \end{aligned}$$

Die asymptotische Güte $\pi_l(h)$ des randomisierten Cramer-von-Mises-Tests φ_{RC_n} aus (4.35) für Beobachtungen mit den Verteilungen $P_{l,n,h}$ in (4.44) ist gegeben durch

$$\pi_l(h) = P(S_l > c_{1-\alpha}). \quad (4.48)$$

Die Abbildung 4.3 zeigen die Gütefunktionen $\pi_1(h), \dots, \pi_4(h)$, die mit Hilfe von Simulationen ermittelt wurden.

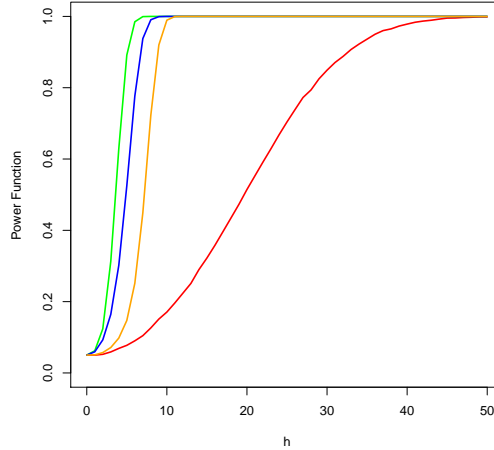


Abbildung 4.3: Verlauf der Gütefunktionen: $\pi_1(h)$ (rot), $\pi_2(h)$ (grün), $\pi_3(h)$ (blau) und $\pi_4(h)$ (orange).

Um den Verlauf der Gütefunktionen zu diskutieren, bemerken wir zunächst, dass die γ_l in (4.45) gerade das Skalarprodukt der Tangente des Modell \dot{l}_0 und der Tangente $\varphi_l(\Phi(t))$ der Folge von Verteilungen $P_{l,n,h}$ in (4.44) sind. Da die Gütefunktion nur von der Projektion von $\varphi_l(\Phi(t))$ auf das orthogonale Komplement des Tangentialraumes abhängt, können wir eine hohe Güte für kleine $|\gamma_l|$ erwarten. Wir wissen bereits $\gamma_{2m} = 0$. Die γ_l wurden für ungerade l numerisch berechnet. Die ersten Werte sind in der folgenden Tabelle dargestellt.

Diese Werte werden durch die Abbildung 4.4 veranschaulicht, die Bilder stellen die Integranden in (4.45) für $l = 1, 2, 3$ und 4 dar. Wenn die entsprechende Funktion ist vor allem dann positiv

4 Anpassungstests

$\gamma_1 = -0.9484$	$\gamma_3 = -0.2407$	$\gamma_5 = -0.1306$	$\gamma_7 = -0.0880$
$\gamma_9 = -0.0657$	$\gamma_{11} = -0.0522$	$\gamma_{13} = -0.0431$	$\gamma_{15} = -0.0366$

Tabelle 4.2: Werte von γ_l für $l = 1, 2, \dots, 15$

oder negativ $|\gamma_l|$ groß ist. Betrachten wir zum Beispiel der Fall $l = 1$. Die rote Kurve entspricht dem Wert $|\gamma_1| = -0.9484$.

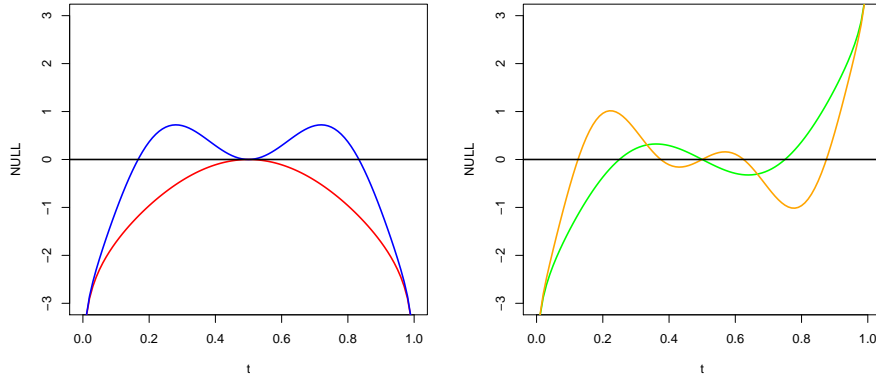


Abbildung 4.4: Verlauf des Integranden. Linkes Bild: $l = 1$ (rot), $l = 3$ (blau). Rechtes Bild: $l = 2$ (grün), $l = 4$ (orange).

Um den schwachen Anstieg der Gütefunktion $\pi_1(h)$ zu erklären, betrachten wir den führenden Term in der Statistik S_1

$$\begin{aligned}
 & \frac{1}{\pi^2}(Z_1 + h(1 - \gamma_1^2))^2 + \frac{1}{(3\pi)^2}(Z_3 - h\gamma_1\gamma_3)^2 \\
 & + \frac{1}{(5\pi)^2}(Z_5 - h\gamma_1\gamma_5)^2 + \frac{1}{(7\pi)^2}(Z_7 - h\gamma_1\gamma_7)^2 + \dots \\
 \approx & 0.1013(Z_1 + 0.01h)^2 + 0.0113(Z_3 + 0.2283h)^2 \\
 & + 0.0041(Z_5 - 0.1239h)^2 + 0.0021(Z_7 - 0.083h)^2 + \dots
 \end{aligned}$$

Wir sehen, dass der Term auf der rechten Seite nur sehr schwach von h abhängt, was die unbefriedigende Güte erklärt. Die größte Gütefunktion ist $\pi_2(h)$. Der starke Anstieg kann durch die Untersuchung der führenden Terme in S_2 erklärt werden, die durch

$$0.1013Z_1^2 + 0.0253(Z_2 + h)^2 + 0.0113Z_3^2 + \dots$$

gegeben sind. Verglichen mit S_1 liegt jetzt eine wesentlich stärkere Abhängigkeit vor. Zusammenfassend können wir sagen, dass im Gaußschen Lokationsmodell der randomisierte Cramer-von-Mises-Test eine geringe Güte hat für die Alternativen in Abbildung 4.2 mit $l = 1$ und eine große Güte hat für Alternativen aus der Abbildung 4.1 mit $l = 2$.

4.3.2 Normalverteilung mit unbekannten μ und σ^2

Falls μ und σ^2 unbekannt sind, dann gilt $P_\theta = N(\mu, \sigma^2)$, $\theta = (\mu, \sigma^2)$, $\mu \in \mathbb{R}$, $\sigma^2 > 0$ und $F_\theta(t) = \Phi\left(\frac{t-\mu}{\sigma}\right)$, sowie

$$\begin{aligned}\dot{F}_\theta(t) &= \left(\frac{\partial}{\partial \mu} \Phi \left(\frac{t - \mu}{\sigma} \right), \frac{\partial}{\partial \sigma^2} \Phi \left(\frac{t - \mu}{\sigma} \right) \right)^T \\ &= \left(-\frac{1}{\sigma} \varphi \left(\frac{t - \mu}{\sigma} \right), -\frac{(t - \mu)}{2\sigma^3} \varphi \left(\frac{t - \mu}{\sigma} \right) \right)^T.\end{aligned}$$

Die Bedingung (4.20) ist wieder erfüllt. Der Maximum-Likelihood-Schätzer für $\theta = (\mu, \sigma^2)$ ist

$$\hat{\theta}_n = (\bar{X}_n, S_n^2)^T \text{ mit } S_n^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X}_n)^2.$$

Wegen $\sqrt{n}(\bar{X}_n - \mu)^2 = o_P(1)$ folgen

$$\sqrt{n} \begin{pmatrix} \bar{X}_n - \mu \\ S_n^2 - \sigma^2 \end{pmatrix} = \frac{1}{\sqrt{n}} \sum_{i=1}^n \begin{pmatrix} X_i - \mu \\ (X_i - \mu)^2 - \sigma^2 \end{pmatrix} + o_P(1)$$

und

$$\mathcal{L} \left(\sqrt{n} \begin{pmatrix} \bar{X}_n - \mu \\ S_n^2 - \sigma^2 \end{pmatrix} \right) \Rightarrow N(0, \Sigma),$$

wobei

$$\Sigma = \begin{pmatrix} \sigma^2 & 0 \\ 0 & 2\sigma^4 \end{pmatrix} = I^{-1}(\theta)$$

ist.

Folglich sind die Bedingungen (4.17), (4.18) und (4.19) erfüllt. Weiterhin ist (4.27) erfüllt. Für den geschätzten empirischen Prozess

$$\hat{G}_n(t) = \sqrt{n} \left(\hat{F}_n(t) - \Phi \left(\frac{t - \bar{X}_n}{S_n} \right) \right)$$

gilt deshalb nach Satz (4.12) $\mathcal{L}(\hat{G}_n) \Rightarrow \mathcal{L}(Z_{(\mu, \sigma^2)})$, wobei $Z_{(\mu, \sigma^2)}$ ein zentrierter Gaußprozess ist, dessen Kovarianzfunktion nach (4.29) gegeben ist durch

$$\begin{aligned}\text{cov}(Z_{(\mu, \sigma^2)}(s), Z_{(\mu, \sigma^2)}(t)) &= \Phi \left(\frac{(s \wedge t) - \mu}{\sigma} \right) - \Phi \left(\frac{s - \mu}{\sigma} \right) \Phi \left(\frac{t - \mu}{\sigma} \right) \\ &\quad - \varphi \left(\frac{s - \mu}{\sigma} \right) \varphi \left(\frac{t - \mu}{\sigma} \right) - \frac{(s - \mu)(t - \mu)}{2\sigma^2} \varphi \left(\frac{s - \mu}{\sigma} \right) \varphi \left(\frac{t - \mu}{\sigma} \right).\end{aligned}$$

Für die Cramer-von-Mises-Statistik

$$\begin{aligned}C_n &= n \int \left(\hat{F}_n(t) - \Phi \left(\frac{t - \bar{X}_n}{S_n} \right) \right)^2 d\hat{F}_n(t) \\ &= \sum_{i=1}^n \left(\frac{i}{n} - \Phi \left((X_{n:i} - \bar{X}_n) / S_n \right) \right)^2\end{aligned} \tag{4.49}$$

erhalten wir aus Folgerung 4.10

$$\mathcal{L}(C_n | P_n^{\otimes n}) \Rightarrow \mathcal{L} \left(\int (Z(t) + A(t) - A_\theta(t))^2 P_\theta(dt) \right).$$

4 Anpassungstests

Der randomisierte geschätzte empirische Prozess $\widehat{BR}_n(t)$ in (4.30) ist gegeben durch

$$\begin{aligned}\widehat{BR}_n(t) = \sqrt{n} \left(\widehat{F}_n(t) - \Phi \left(\frac{t - \bar{X}_n}{S_n} \right) \right) - \frac{1}{\sqrt{n}} \varphi \left(\frac{t - \bar{X}_n}{S_n} \right) V_{1,n} \\ - \frac{(t - \bar{X}_n)}{\sqrt{2n} S_n} \varphi \left(\frac{t - \bar{X}_n}{S_n} \right) V_{2,n},\end{aligned}$$

wobei $V_{1,n}, V_{2,n}$ unabhängig, standardnormal und unabhängig von X_1, \dots, X_n sind. Die randomisierte Cramer-von-Mises-Statistik ist gegeben durch

$$\begin{aligned}RC_n = \sum_{i=1}^n \left(\frac{i}{n} - \Phi \left((X_{n:i} - \bar{X}_n) / S_n \right) \right. \\ \left. - \frac{1}{\sqrt{n}} \varphi \left((X_{n:i} - \bar{X}_n) / S_n \right) V_{1,n} - \frac{(X_{n:i} - \bar{X}_n)}{\sqrt{2n} S_n} \varphi \left((X_{n:i} - \bar{X}_n) / S_n \right) V_{2,n} \right)^2.\end{aligned}\quad (4.50)$$

4.3.3 Exponentialverteilung mit unbekanntem λ

Seien

$$G(t) = I_{[0,\infty)}(t)(1 - \exp\{-t\}) \quad \text{und} \quad g(t) = I_{[0,\infty)}(t) \exp\{-t\},$$

die Verteilungsfunktion und die Dichte der Standardexponentialverteilung. Das parametrisierte Modell der Exponentialverteilung ist gegeben durch

$$G_\theta(t) = I_{[0,\infty)}(t)(1 - \exp\{-\theta t\}) \quad \text{und} \quad \dot{G}_\theta(t) = I_{[0,\infty)}(t)(t \exp\{-\theta t\}), \quad \theta > 0.$$

Die Bedingung (4.20) ist erfüllt. Der Maximum-Likelihood-Schätzer ist $\hat{\theta}_n = 1/\bar{X}_n$ und es gilt

$$\sqrt{n} \left(\frac{1}{\bar{X}_n} - \theta \right) = -\frac{\theta}{\bar{X}_n} \frac{1}{\sqrt{n}} \sum_{i=1}^n (X_i - \frac{1}{\theta}).$$

Wegen $V(X_i) = \frac{1}{\theta^2}$ sehen wir, dass die Bedingungen (4.17), (4.18) und (4.19) erfüllt sind mit $h_\theta(x) = \theta^2(1/\theta - x)$ und $J(\theta) = I(\theta) = 1/\theta^2$. Weiterhin ist (4.27) erfüllt. Die Cramer-von-Mises-Statistik lautet

$$\begin{aligned}C_n &= n \int \left(\widehat{F}_n(t) - G_{1/\bar{X}_n}(t) \right)^2 d\widehat{F}_n(t) \\ &= \sum_{i=1}^n \left(\frac{i}{n} - (1 - \exp\{-X_{n:i}/\bar{X}_n\}) \right)^2.\end{aligned}$$

Die randomisierte Cramer-von-Mises-Statistik ist gegeben durch

$$RC_n = \sum_{i=1}^n \frac{i}{n} - (1 - \exp\{-X_{n:i}/\bar{X}_n\}) \quad (4.51)$$

$$- \frac{V_n}{\sqrt{n}} \frac{X_{n:i}}{\bar{X}_n} \exp\{-X_{n:i}/\bar{X}_n\}^2. \quad (4.52)$$

4.4 Computersimulationen

4.4.1 Niveau der Anpassungstests

Monte-Carlo-Simulationen sind ein systematisches Hilfsmittel, um zu überprüfen, ob asymptotische α -Tests das Niveau bereits bei kleinen oder mittleren Stichprobenumfängen einhalten. Gleichzeitig geben solche Simulationen Einblick in die Struktur der Gütefunktion für fest gewählte Alternativen. Umfangreiche Simulationen für die Überprüfung des Cramer-von-Mises-Tests und anderer Anpassungstests hinsichtlich des Einhaltens des vorgegebenen Niveaus wurden von vielen Autoren durchgeführt. Wir verweisen auf Stephens (1974) und Stephens (1976). Entsprechende Simulationen haben wir für die randomisierten Tests durchgeführt.

1. Normalverteilung μ und σ^2 unbekannt

Wir verwenden die Bezeichnungen aus dem Abschnitt 4.3.1 und setzen $\mu = 0$ und $\sigma^2 = 1$ voraus. Die randomisierte Cramer-von-Mises-Statistik (4.50) ist gegeben durch

$$RC_n = \sum_{i=1}^n \left(\frac{i}{n} - \Phi((X_{n:i} - \bar{X}_n)/S_n) - \frac{1}{\sqrt{n}} \varphi((X_{n:i} - \bar{X}_n)/S_n) V_{1,n} - \frac{(X_{n:i} - \bar{X}_n)}{\sqrt{2n}S_n} \varphi((X_{n:i} - \bar{X}_n)/S_n) V_{2,n} \right)^2. \quad (4.53)$$

2. Exponentialverteilung mit unbekanntem λ

Wir verwenden die Bezeichnungen aus dem Abschnitt 4.3.3 und setzen $\lambda = 1$ voraus. Die randomisierte Cramer-von-Mises-Statistik in (4.52) ist gegeben durch

$$RC_n = \sum_{i=1}^n \left(\frac{i}{n} - (1 - \exp\{-X_{n:i}/\bar{X}_n\}) - \frac{V_n}{\sqrt{n}} \frac{X_{n:i}}{\bar{X}_n} \exp\{-X_{n:i}/\bar{X}_n\} \right)^2. \quad (4.54)$$

Die Simulationen wurden in folgenden Schritten durchgeführt. Durch F bezeichnen wir entweder die Standardnormalverteilungsfunktion oder die Verteilungsfunktion der Exponentialverteilung mit dem Parameter $\lambda = 1$. T_n ist eine der beiden Statistiken in (4.53) oder (4.54). Die Simulationen wurden mit Hilfe des Softwarepaketes R ausgeführt.

1. Für $n = 20, 50, 100$ und 1000 erzeugen wir X_1, \dots, X_n mit der Verteilungsfunktion F .
2. Wir berechnen den Maximum-Likelihood-Schätzer $\hat{\theta}_n$.
3. Wir berechnen den Wert T_n der Teststatistik.
4. Wir führen den Test

$$\varphi_n = \begin{cases} 1 & , \quad T_n > c_{n,1-\alpha} \\ 0 & , \quad \text{sonst} \end{cases}$$

durch, wobei die $c_{n,1-\alpha}$ die exakten $1 - \alpha$ -Quantile des Cramer-von-Mises-Tests für eine einfache Nullhypothese sind. Die konkreten Werte für $c_{n,1-\alpha}$ sind aus Shorack und Wellner (1986).

4 Anpassungstests

5. Die Schritte 1.-4. werden N -mal wiederholt und das tatsächliche Konfidenzniveau wird geschätzt durch

$$\hat{\alpha}_{T_n} = \frac{\text{Anzahl der Ablehnungen von } H_0}{N}.$$

In unseren Simulationen haben wir $N = 10000$ verwendet.

In der Tabelle 4.3 bezeichnet $\alpha(\text{CMR})$ das geschätzte Niveau des randomisierten Cramer-von-Mises-Tests.

$\alpha(\text{CMR})$: Normalverteilung, μ, σ^2 unbekannt				
α	$n = 20$	$n = 50$	$n = 100$	$n = 1000$
0.01	0.013	0.010	0.010	0.010
0.05	0.060	0.052	0.053	0.047
0.1	0.114	0.104	0.107	0.099
$\alpha(\text{CMR})$: Exponentialverteilung, λ unbekannt				
0.01	0.012	0.012	0.012	0.009
0.05	0.055	0.051	0.053	0.049
0.1	0.108	0.104	0.105	0.099

Tabelle 4.3: Geschätztes Niveau des randomisierten Cramer-von-Mises-Tests.

Aus der Tabelle 4.3 können wir ablesen, dass sich die randomisierte Statistik RC_n bereits für kleine n sehr ähnlich verhält wie die Cramer-von-Mises-Statistik in der klassischen Situation einer einfachen Nullhypothese, das tatsächliche Niveau kommt dem vorgegebenem α bereits für kleine n sehr nahe.

4.4.2 Die asymptotische Güte unter speziellen Alternativen

Seien X_1, \dots, X_n u.i.v. mit einer Normalverteilung mit den Parametern μ und σ^2 . Sei $X_{n:1} \leq \dots \leq X_{n:n}$ die Ordnungsstatistik. V bezeichnet die Kovarianzmatrix von $(X_{n:1}, \dots, X_{n:n})$ und m sei der Erwartungswertvektor. Wir setzen

$$a^T = (a_1, \dots, a_n) = \frac{m^T V^{-1}}{(m^T V^{-1} V^{-1} m)^{1/2}}.$$

Die Shapiro-Wilk-Statistik ist definiert durch

$$W_n = \frac{\sum_{i=1}^n a_i X_{n:i}}{\sum_{i=1}^n (X_i - \bar{X}_n)^2}.$$

Nach Shapiro und Francia (1972) kann diese Statistik für $n > 50$ durch folgende Statistik approximiert werden

$$W_n^* = \frac{\sum_{i=1}^n b_i X_{n:i}}{\sum_{i=1}^n (X_i - \bar{X}_n)^2},$$

$$b = (b_1, \dots, b_n) = \frac{1}{\sqrt{m^T m}} m.$$

Sei X eine Zufallsvariable mit $EX^8 < \infty$. Dann sind Schiefe und Exzess definiert durch

$$\beta_1 = \frac{1}{\sigma^3} E(X - \mu)^3 \quad \text{und} \quad \beta_2 = \frac{1}{\sigma^4} E(X - \mu)^4 - 3,$$

wobei $\mu = EX$, $\sigma^2 = V(X)$ gesetzt wurden. Falls X normalverteilt ist, dann gilt $\beta_1 = \beta_2 = 0$, so dass die Bowman-Shenton-Statistik

$$\begin{aligned} BS_n &= \frac{n}{6}(\hat{\beta}_{1,n})^2 + \frac{n}{24}(\hat{\beta}_{2,n})^2, \\ \hat{\beta}_{1,n} &= \frac{1}{S_n^3} \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X}_n)^3 \quad \hat{\beta}_{2,n} = \frac{1}{S_n^4} \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X}_n)^4 - 3 \end{aligned} \quad (4.55)$$

eine mögliche Schiefe und einen möglichen Exzess registriert und für das Testen auf Normalverteilung verwendet werden kann. Bowman und Shenton (1975) bewiesen, dass $\mathcal{L}(BS_n)$ schwach gegen eine Normalverteilung mit zwei Freiheitsgraden strebt. Der Bowman-Shenton-Test lehnt die Nullhypothese (Daten sind normalverteilt) für große Werte von BS_n ab. In D'Agostino und Stephens (1986) und anderen Veröffentlichungen wurde bemerkt, dass $\mathcal{L}(BS_n)$ nur sehr langsam konvergiert und es wurden deshalb Transformationen von BS_n vorgeschlagen. Wir verwenden nicht diesen Weg. Stattdessen bestimmen wir die Quantile von BS_n durch Simulationen.

Um den Shapiro-Wilk-Test und den Bowman-Shenton-Test mit den von uns entwickelten randomisierten Tests zu vergleichen, betrachten wir zwei Typen von parametrisierten Alternativen.

1. Abweichungen von der Normalverteilung in Richtung Schiefe

X möge eine Standardnormalverteilung besitzen und Y möge eine Exponentialverteilung mit dem Parameter λ besitzen. Wir setzen $P_\lambda = \mathcal{L}(X + \lambda Y)$, $\lambda > 0$. P_0 ist die Standardnormalverteilung und λ kontrolliert die Abweichung von einer Normalverteilung.

2. Abweichung von der Normalverteilung in Richtung schwerer Flanken

X möge eine Standardnormalverteilung besitzen und C eine Cauchy-Verteilung. Sei $Q_\lambda = \mathcal{L}(X + \lambda C)$, $\lambda > 0$. Wieder ist Q_0 die Standardnormalverteilung und λ kontrolliert die Abweichung von einer Normalverteilung, jetzt aber in Richtung schwerer Flanken.

Für die beiden Alternativen stellen wir in der Abbildungen 4.5 und 4.6 die Gütefunktionen des Cramer-von-Mises-Tests, des Shapiro-Wilk-Tests und Bowman-Shenton-Tests dar.

Die Ergebnisse stimmen mit den Funktionen π_i in (4.48) und dem Verlauf der Dichtefunktionen in den Abbildungen 4.1 und 4.2 überein. Die Gütefunktion π_1 hat einen schwachen Anstieg.

Die Dichte $p_{1,\eta}$ in Abbildung 4.2 ist schief. Es ist deshalb nicht überraschend, dass der randomisierte Cramer-von-Mises-Test eine geringe Güte für schiefe Alternativen hat. Dagegen hat die Gütefunktion π_2 einen starken Anstieg und die Dichten $p_{2,\eta}$ sind breiter als die Dichten der Normalverteilung, so dass eine gute Güte für den Cramer-von-Mises-Test erwartet werden kann. Diese Vermutung wird durch die obigen Simulationsergebnisse bestätigt.

4 Anpassungstests

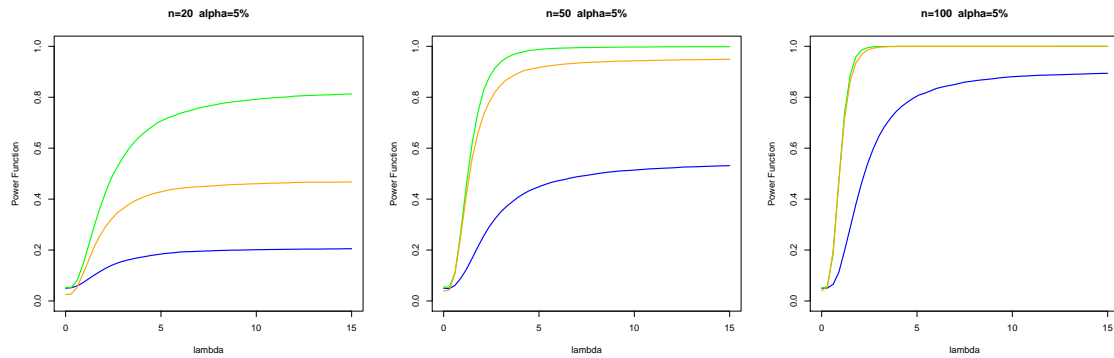


Abbildung 4.5: Verlauf der Gütefunktionen des Cramer-von-Mises-Tests (blau), Shapiro-Wilk-Tests (grün) und Bowman-Shenton-Tests (orange) unter der ersten Alternative

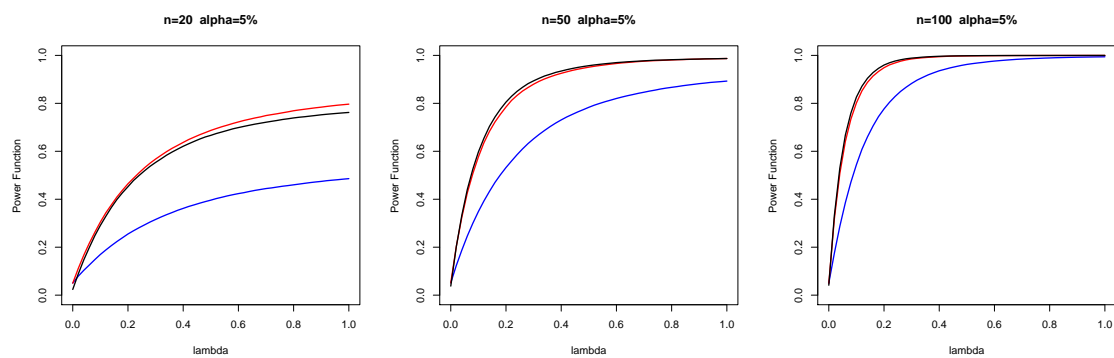


Abbildung 4.6: Verlauf der Gütefunktionen des Cramer-von-Mises-Tests (blau), Shapiro-Wilk-Tests (rot) und Bowman-Shenton-Tests (schwarz) unter der zweiten Alternative

5 Anhang

5.1 Anhang I: Die Lemmata von LeCam

In diesem Anhang stellen wir Resultate aus der asymptotischen Entscheidungstheorie zusammen, wobei wir für die Beweise größtenteils auf die Literatur verweisen.

Für einen messbaren Raum $(\mathcal{X}, \mathfrak{A})$ und eine Verteilung P auf $(\mathcal{X}, \mathfrak{A})$ bezeichnen wir durch $L_2(P)$ den Raum der quadratisch integrierbaren Funktionen $a : \mathcal{X} \rightarrow \mathbb{R}$. Weiter sei

$$L_2^0(P) = \left\{ a : a \in L_2(P), \int a dP = 0 \right\}.$$

Für die Nullhypothese $H_0 : P$ untersuchen wir die Folge von Hypothesen $H_A : P_n$ für die die Folge $\{P_n^{\otimes n}\}$ benachbart zur Folge $\{P^{\otimes n}\}$ ($P_n^{\otimes n} \triangleleft P^{\otimes n}$) ist, d. h.

$$\lim_{n \rightarrow \infty} P_n^{\otimes n}(A_n) = 0 \quad \Rightarrow \quad \lim_{n \rightarrow \infty} P_n^{\otimes n}(A_n) = 0. \quad (5.1)$$

Um eine Folge $\{P_n\}$ zu konstruieren, die der Bedingung (5.1) genügt, setzen wir $P_n \ll P$ voraus. Seien

$$g_n = \frac{dP_n}{dP} \quad \text{und} \quad a_n = 2\sqrt{n}(\sqrt{g_n} - 1). \quad (5.2)$$

Wir nehmen an, dass die Folge $a_n \in L_2(P)$ gegen $a \in L_2(P)$ strebt, d. h.

$$\lim_{n \rightarrow \infty} \int (a_n(x) - a(x))^2 P(dx) = 0. \quad (5.3)$$

Das nächste Lemma zeigt, dass dann die Bedingung (5.1) erfüllt ist. Darüber hinaus liefert dieses Lemma eine Entwicklung der Log-Likelihoodfunktion.

Lemma 5.1 Seien $X_{n,1}, \dots, X_{n,n}$ die Projektionen von \mathcal{X}^n auf \mathcal{X} und es sei $P_n \ll P$. Sei $g_n = dP_n/dP$ und $a_n = 2\sqrt{n}(\sqrt{g_n} - 1)$ möge (4.4) erfüllen für ein $a \in L_2^0(P)$. Dann gelten

$$\ln L_n = \frac{1}{\sqrt{n}} \sum_{i=1}^n a(X_{n,i}) - \frac{1}{2} \int_0^1 a^2(x) P(dx) + o_{P^{\otimes n}}(1), \quad (5.4)$$

$$\begin{aligned} L_n &= \frac{dP_n^{\otimes n}}{dP^{\otimes n}}(X_{n,1}, \dots, X_{n,n}), \\ P_n^{\otimes n} &\triangleleft P^{\otimes n}. \end{aligned} \quad (5.5)$$

Die Aussage (5.4) ist ein Spezialfall von Satz 3 in Shorack und Wellner (1986), S. 154. Die Tatsache, dass (5.4) die Aussage (5.5) impliziert, ist der Inhalt des ersten Lemmas von LeCam, siehe Exercise 2, S. 157 in Shorack und Wellner (1986).

Für $r_n = a_n - a$ ergibt die Definition von a_n

$$\sqrt{g_n} - 1 = \frac{1}{2\sqrt{n}} a + \frac{1}{2\sqrt{n}} r_n,$$

wobei $\int r_n^2 dP \rightarrow 0$ gilt, falls (4.4) erfüllt ist. Deshalb nennen wir a die *Tangente* der Folge $\{P_n\}$. Das nächste Lemma zeigt, dass die Menge aller Tangenten von Folgen gerade mit $L_2^0(P)$ übereinstimmt.

Lemma 5.2 Falls die Bedingung (4.4) erfüllt ist, dann gilt $a \in L_2^0(P)$. Umgekehrt gibt es für jedes $a \in L_2^0(P)$ ein $\varepsilon > 0$ und eine Kurve (Q_η) , $-\varepsilon < \eta < \varepsilon$ derart, dass gilt $Q_0 = P$ und

$$\lim_{\eta \rightarrow 0} \frac{1}{\eta^2} \int \left(\sqrt{\frac{dQ_\eta}{dP}} - 1 - \frac{\eta}{2} a \right)^2 dP < \infty. \quad (5.6)$$

Insbesondere ist also $P_n = Q_{1/\sqrt{n}}$ eine Folge von Verteilungen mit $P_n \ll P$ und der Tangente a .

Für den Beweis verweisen wir auf Janssen (2004). Ein zentraler Begriff der asymptotischen Statistik ist der Begriff der L_2 -Differenzierbarkeit eines Modells, der oben bereits implizit verwendet wurde. Sei $\mathcal{P} = (P_\theta)_{\theta \in \Theta}$, $\Theta \subseteq \mathbb{R}_d$ eine parametrisierte Verteilungsfamilie auf $(\mathcal{X}, \mathfrak{A})$. Ohne wesentliche Einschränkung nehmen wir an, dass je zwei Verteilungen $P_{\theta_1}, P_{\theta_2}$ gegenseitig absolutstetig sind.

Definition 5.3 Die Familie $(P_\theta)_{\theta \in \Theta}$ heißt in dem inneren Punkt $\theta \in \Theta^0$ L_2 -differenzierbar, falls es eine vektorwertige messbare Funktion $\dot{L}_\theta : \mathcal{X} \rightarrow \mathbb{R}_d$ gibt, L_2 -Ableitung im Punkt θ genannt, mit den Eigenschaften

$$\begin{aligned} \int \|\dot{L}_\theta\|^2 dP_\theta &< \infty, \\ \int \left(\sqrt{\frac{dP_{\theta+u}}{dP_\theta}} - 1 - \frac{1}{2} u^T \dot{L}_\theta \right)^2 dP_\theta &= o(\|u\|^2). \end{aligned}$$

Aus der obigen Definition erhalten wir

$$\sqrt{\frac{dP_{\theta+u}}{dP_\theta}} = 1 + \frac{1}{2} u^T \dot{L}_\theta + R(u), \quad \int R^2(u) dP_\theta = o(\|u\|^2).$$

Das ergibt

$$\begin{aligned} 1 &= \int \left(1 + \frac{1}{2} u^T \dot{L}_\theta + R(u) \right)^2 dP_\theta \\ &= \int \left(1 + \frac{1}{4} (u^T \dot{L}_\theta)^2 + R^2(u) + u^T \dot{L}_\theta + u^T \dot{L}_\theta R(u) + 2R(u) \right) dP_\theta. \end{aligned}$$

Weiterhin sind

$$\int (u^T \dot{L}_\theta)^2 dP_\theta = o(\|u\|^2), \quad \int u^T \dot{L}_\theta R(u) = o(\|u\|^2), \quad \int R(u) dP_\theta = o(\|u\|).$$

Das ergibt

$$\int u^T \dot{L}_\theta dP_\theta = o(\|u\|).$$

Das ist nur möglich, wenn

$$\int \dot{L}_\theta dP_\theta = 0 \quad (5.7)$$

gilt. Wir verstehen unter dem *Tangentialraum* $\mathcal{T}(\theta)$ im Punkt θ eines im Punkt θ L_2 -differenzierbaren Modells die Menge der Funktionen

$$\mathcal{T}(\theta) = \{h^T \dot{L}_\theta, h \in \mathbb{R}_d\}.$$

Aus (5.7) folgt dass $\mathcal{T}(\theta)$ ein abgeschlossener Unterraum von $L_2^0(P_\theta)$ ist, der nach Lemma 5.2 mit dem Raum aller Tangenten von einparametrischen L_2 -differenzierbaren Kurven $(Q_\eta)_{-\varepsilon < \eta < \varepsilon}$ von Verteilungen durch P_θ , d. h. $Q_0 = P_\theta$, zusammenfällt. Es ist leicht zu sehen, dass der Tangentialraum gerade aus den Tangenten von differenzierbaren Kurven besteht, die im Modell liegen, d. h. $(Q_\eta)_{-\varepsilon < \eta < \varepsilon} \subseteq \mathcal{P}$. Sei

$$I(\theta) = \int \dot{L}_\theta \dot{L}_\theta^T dP_\theta$$

die Fisher-Informationsmatrix.

Wir berechnen jetzt die Projektion auf den Tangentialraum.

Lemma 5.4 *Bezeichnet $\Pi_\theta a$ die Projektion von $a \in L_2^0(P_\theta)$ auf $\mathcal{T}(\theta)$ und ist die Fisher-Informationsmatrix regulär, so gilt*

$$\Pi_\theta a = c^T(\theta) I^{-1}(\theta) \dot{L}_\theta \text{ und } c(\theta) = \int a \dot{L}_\theta dP_\theta.$$

Die Projektion $\Pi_\theta^\perp a$ von a auf das orthogonale Komplement von $\mathcal{T}(\theta)$ ist gegeben durch

$$\Pi_\theta^\perp a = a - c^T(\theta) I^{-1}(\theta) \dot{L}_\theta.$$

Beweis. $\Pi_\theta a$ ist durch folgende Bedingungen charakterisiert: $\Pi_\theta a \in \mathcal{T}(\theta)$ und $a - \Pi_\theta a \perp \mathcal{T}(\theta)$. Setzen wir $h^T = c^T(\theta) I^{-1}(\theta)$, so folgt $\Pi_\theta a \in \mathcal{T}(\theta)$ aus der Definition des Tangentialraumes. Die Funktionen aus $\mathcal{T}(\theta)$ haben die Gestalt $h^T \dot{L}_\theta$ und es gilt

$$\begin{aligned} \int h^T \dot{L}_\theta (a - c^T(\theta) I^{-1}(\theta) \dot{L}_\theta) dP_\theta &= \int h^T \dot{L}_\theta (a - \dot{L}_\theta^T I^{-1}(\theta) c(\theta)) dP_\theta \\ &= h^T \left(\int a \dot{L}_\theta dP_\theta - c(\theta) \right) = 0. \end{aligned}$$

■

Wir betrachten jetzt lineare Statistiken der folgenden Form

$$T_n = \frac{1}{\sqrt{n}} \sum_{i=1}^n b_n(X_i) + o_{P^{\otimes n}}(1). \quad (5.8)$$

Proposition 5.5 *Es seien X_1, \dots, X_n u.i.v. mit der Verteilung $P_n \ll P$ und es sei die Bedingung (4.4) erfüllt mit $g_n = dP_n/dP$ und a_n von (5.2). Falls $b_n \in L_2^0(P)$ die Bedingung*

$$\lim_{n \rightarrow \infty} \int (b_n - b)^2 dP = 0 \quad (5.9)$$

für ein $b \in L_2(P)$ erfüllt, dann gilt

$$\mathcal{L}(T_n | P_n^{\otimes n}) \Rightarrow N \left(\int ab dP, \int b^2 dP \right).$$

Beweis. Die Bedingung (5.9) und das Lemma von Slutsky implizieren, dass man b_n durch b in (5.8) ersetzen kann. Das ergibt zusammen mit der Darstellung (5.4) im Lemma 5.1

$$\begin{pmatrix} T_n \\ \ln L_n \end{pmatrix} = \frac{1}{\sqrt{n}} \sum_{i=1}^n \begin{pmatrix} b(X_i) \\ a(X_i) \end{pmatrix} + \begin{pmatrix} 0 \\ -\frac{1}{2} \int_0^1 a^2(x) P(dx) \end{pmatrix} + o_{P^{\otimes n}}(1).$$

Die Anwendung des zentralen Grenzwertsatzes ergibt

$$\mathcal{L}((T_n, \ln L_n)^T | P^{\otimes n}) \Rightarrow N \left(\begin{pmatrix} 0 \\ -\frac{1}{2} \int_0^1 a^2(x) P(dx) \end{pmatrix}, \begin{pmatrix} \int_0^1 b^2 dP & \int ab dP \\ \int ab dP & \int_0^1 a^2 dP \end{pmatrix} \right).$$

Diese Aussage und das dritte Lemma von LeCam, siehe Satz 4 in Shorack und Wellner (1986), S. 154, ergeben (4.7). ■

5.2 Anhang II: Schwache Konvergenz im Skorochod-Raum

Sei $x \in \mathbb{D}[-\infty, \infty]$. Weil die Koordinatenabbildungen $x \mapsto x(t)$ stetig und messbar sind, folgt aus der schwachen Konvergenz der Verteilungen, dass für alle Gitter $-\infty \leq t_0 < t_1 < \dots < t_N \leq \infty$ für $n \rightarrow \infty$ gilt

$$\mathcal{L}((Z_n(t_0), \dots, Z_n(t_N))|P_n) \Rightarrow \mathcal{L}((Z(t_0), \dots, Z(t_N))|P_n).$$

Dann sprechen wir von der Konvergenz aller endlich dimensional Verteilungen und bezeichnen dies durch

$$\mathcal{L}(Z_n|P) \Rightarrow^{fidi} \mathcal{L}(Z|P). \quad (5.10)$$

Die umgekehrte Aussage ist nur unter Zusatzvoraussetzungen möglich. Hierzu gibt es verschiedene Varianten. Wir verwenden hier den im Buch von Pollard verwendeten Zugang. Das nächste Lemma folgt aus Satz 3 auf S. 92, Pollard (1984).

Lemma 5.6 *Es seien Z, Z_1, Z_2, \dots Zufallselemente, die auf $(\Omega, \mathfrak{F}, Q)$ definiert sind und Werte in $\mathbb{D}[-\infty, \infty]$ haben, wobei dieser Raum mit der Supremumsmetrik ausgestattet ist und die σ -Algebra $\mathfrak{B}_{\mathbb{D}}$ von Teilmengen von \mathbb{D} durch die Projektionen erzeugt wird. Dann gilt $\mathcal{L}(Z_n) \Rightarrow \mathcal{L}(Z)$ dann und nur dann, wenn die endlich dimensional Verteilungen von Z_n schwach gegen die entsprechenden endlich dimensional Verteilungen von Z konvergieren und für jedes $\varepsilon > 0$ und jedes $\delta > 0$ ein Gitter $-\infty = t_0 < t_1 < \dots < t_N = \infty$ existiert mit*

$$\limsup_{n \rightarrow \infty} Q \left(\max_{0 \leq i < N-1} \sup_{t_i \leq t < t_{i+1}} |Z_n(t) - Z_n(t_i)| > \delta \right) < \varepsilon. \quad (5.11)$$

Lemma 5.7 *Es seien Y, Y_1, Y_2, \dots zufällige Elemente von $\mathbb{D}[-\infty, \infty]$. Die endlich dimensional Verteilungen von Y_n mögen schwach gegen die entsprechenden endlich dimensional Verteilungen von Y konvergieren. Falls $\mathcal{L}(Y_n) \triangleleft \mathcal{L}(Z_n)$ gilt und Z_n der Bedingung (5.11) genügt, dann folgt $\mathcal{L}(Y_n) \Rightarrow \mathcal{L}(Y)$.*

Beweis. Nach der Jurečková-Charakterisierung der Benachbarkeit, siehe Lemma 1, S. 157 in Shorack und Wellner (1986), gibt es für jedes $\eta > 0$ ein $\varepsilon > 0$ so dass (5.11) impliziert

$$\limsup_{n \rightarrow \infty} Q \left(\max_{0 \leq i < N-1} \sup_{t_i \leq t < t_{i+1}} |Y_n(t) - Y_n(t_i)| > \delta \right) < \eta.$$

Die Aussage folgt dann aus dem vorigen Lemma. ■

Lemma 5.8 *Es seien Z, Z_1, Z_2, \dots zufällige Elemente mit Werten in $\mathbb{D}[-\infty, \infty]$, wobei Z stetig ist. Die Zufallsvariablen $X_{n,1}, \dots, X_{n,n}$ seien unabhängig und identisch verteilt mit Verteilung P_n und*

$$\hat{F}_n(t) = \frac{1}{n} \sum_{i=1}^n I_{(-\infty, t]}(X_{n,i}).$$

Es gelte $P_n^{\otimes n} \triangleleft P^{\otimes n}$ und $\mathcal{L}(Z_n) \Rightarrow \mathcal{L}(Z)$. Dann folgt

$$\mathcal{L} \left(\int Z_n^2(t) d\hat{F}_n(t) \right) \Rightarrow \mathcal{L} \left(\int Z^2(t) P(dt) \right).$$

Beweis. Seien X_1, \dots, X_n unabhängig und identisch verteilt mit der Verteilung P . Dann folgt mit $F(t) = P((-\infty, t])$ aus dem Satz von Glivenko-Cantelli für jedes $\varepsilon > 0$

$$\lim_{n \rightarrow \infty} P^{\otimes n}(|\sup_t \frac{1}{n} \sum_{i=1}^n I_{(-\infty, t]}(X_i) - F(t)| > \varepsilon) = 0.$$

Die Voraussetzung $P_n^{\otimes n} \triangleleft P^{\otimes n}$ ergibt dann für jedes $\varepsilon > 0$

$$\lim_{n \rightarrow \infty} P_n^{\otimes n}(|\sup_t |\hat{F}_n(t) - F(t)| > \varepsilon) = 0. \quad (5.12)$$

Seien $W_n = Z_n^2$, $W = Z^2$. Für jede Zerlegung $\mathcal{P} = \{t_0, \dots, t_N\}$, $-\infty = t_0 < t_1 < \dots < t_N = \infty$ bezeichnen wir durch $W_{n, \mathcal{P}}$ und $W_{\mathcal{P}}$ die stückweise konstanten Prozesse, die in $[t_i, t_{i+1})$ die Werte $W_n(t_i)$ und $W(t_i)$ haben. Die Konvergenz der endlich dimensionalen Verteilungen und (5.12) ergeben

$$\mathcal{L}\left(\int W_{n, \mathcal{P}} d\hat{F}_n\right) \Rightarrow \mathcal{L}\left(\int W_{\mathcal{P}} dP\right). \quad (5.13)$$

Der stetige Abbildungssatz und $\mathcal{L}(Z_n) \Rightarrow \mathcal{L}(Z)$ implizieren $\mathcal{L}(W_n) \Rightarrow \mathcal{L}(W)$. Folglich gibt es nach dem Lemma 5.6 für jedes m eine Zerlegung \mathcal{P}_m , so dass für alle n gilt

$$\left|\int W_n d\hat{F}_n - \int W_{n, \mathcal{P}_m} d\hat{F}_n\right| < \frac{1}{m} \quad \text{und} \quad \left|\int W dP - \int W_{\mathcal{P}_m} dP\right| < \frac{1}{m}.$$

Die Aussage (5.13) impliziert für jede Lipschitz-stetige Funktion φ mit der Lipschitzkonstanten L

$$\limsup_{n \rightarrow \infty} \left|E\varphi\left(\int W_n d\hat{F}_n\right) - E\varphi\left(\int W dF\right)\right| \leq \frac{2L}{m}.$$

Für $m \rightarrow \infty$ ergibt sich die Behauptung. ■

5.3 Anhang III: Brownsche Brücke und Kac-Siegiert-Zerlegung

Die so genannte Hauptkomponentenzerlegung stochastischer Prozesse spielt für die Theorie der Anpassungstests eine entscheidende Rolle. Es sei $X(t), 0 \leq t \leq 1$ ein messbarer stochastischer Prozess mit endlichen zweiten Momenten, für den gilt $EX(t) = 0, 0 \leq t \leq 1$. Durch

$$K_X(s, t) = \text{cov}(X(s), X(t))$$

bezeichnen wir die Kovarianzfunktion. Wir setzen voraus

$$\int_0^1 \left(\int_0^1 K_X^2(s, t) ds\right) dt < \infty \quad (5.14)$$

und betrachten K_X als linearen Operator im Hilbertraum der quadratisch integrierbaren Funktionen $L_2[0, 1]$, der durch

$$(K_X \varphi)(s) := \int_0^1 K_X(s, t) \varphi(t) dt$$

definiert ist. Eine Zahl $\lambda \neq 0$ heißt Eigenwert zur Eigenfunktion $\varphi \neq 0$, falls

$$K_X \varphi = \lambda \varphi \quad (5.15)$$

gilt. Aus der Tatsache, dass $K_X(s, t)$ positiv semidefinit ist, folgt dass jeder Eigenwert nichtnegativ ist. Weiterhin folgt aus der Bedingung (5.14), dass der lineare Teilraum der Funktionen, die zum gleichen Eigenwert gehören, stets endlich dimensional ist. Diese Dimension heißt die Vielfachheit des entsprechenden Eigenwertes. Weiterhin lässt sich leicht zeigen, dass Eigenfunktionen zu verschiedenen Eigenwerten stets senkrecht aufeinander stehen. Wir bezeichnen durch $\lambda_0 \geq \lambda_1 \geq \dots$ die Folge der Eigenwerte und durch $\varphi_1, \varphi_2, \dots$ die Folge der zugehörigen Eigenfunktionen, wobei die λ_i entsprechend einer eventuell vorliegenden mehrfachen Vielfachheit auch mehrfach aufgeführt sind und die φ_i aus dem gleichen Eigenraum orthonormal gewählt sind. Der Beweis des folgenden Satzes kann in Shorack und Wellner (1986) nachgelesen werden.

Satz 5.9 *Es sei die Voraussetzung (5.14) erfüllt. Dann haben die Zufallsvariablen*

$$V_i = \int_0^1 X(t)\varphi_i(t)dt$$

den Erwartungswert Null und die Varianz λ_i . Sie sind paarweise unkorreliert und es gilt die Kac-Siebert-Zerlegung

$$X(t) = \sum_{i=0}^{\infty} V_i \varphi_i(t), \quad (5.16)$$

wobei die Konvergenz der Reihe in folgendem Sinne zu verstehen ist

$$\lim_{n \rightarrow \infty} E \int_0^1 (X(t) - \sum_{i=0}^n V_i \varphi_i(t))^2 dt = 0.$$

Ist zusätzlich X ein Gaußprozess, so sind die V_0, V_1, \dots unabhängig.

Die Zerlegung (5.16) ist strukturell sehr einfach und für viele Fragen der Statistik sehr nützlich. Das Problem ist aber, dass das Eigenwertproblem (5.15) in der Regel nicht geschlossen lösbar ist. Glücklicherweise sind aber die Spezialfälle des Wienerprozesses und der Brownschen Brücke geschlossen behandelbar.

Beispiel 5.10 *Es sei $B(t), 0 \leq t \leq 1$ die Standard-Brownsche Brücke. Dann gilt für die Kovarianzfunktion $K_B(s, t) = s \wedge t - st$ und das Eigenwertproblem lautet*

$$\int_0^1 (s \wedge t - st)\varphi(t)dt = \lambda\varphi(s).$$

Durch Bilden der zweiten Ableitung nach s lässt sich dieses Problem in ein Randwertproblem für eine Differentialgleichung zweiter Ordnung mit konstanten Koeffizienten überführen. Auf diesem Wege ergibt sich, siehe Shorack und Wellner (1986)

$$\lambda_k = k^2 \pi^2 \quad \text{und} \quad \varphi_k(t) = \sqrt{2} \sin(k\pi t), \quad k = 1, 2, \dots$$

Die φ_k bilden ein vollständiges Orthonormalsystem und es gilt

$$B(t) = \sum_{k=1}^{\infty} Z_k \frac{\sqrt{2} \sin(k\pi t)}{k\pi}, \quad (5.17)$$

wobei die

$$Z_k = \int_0^1 B(t)\varphi_k(t)dt$$

für $k = 1, 2, \dots$ unabhängig und standardnormal sind. Es gilt dann für jede quadratisch integrierbare Funktion $b(t), 0 \leq t \leq 1$

$$B(t) + b(t) = \sum_{k=0}^{\infty} (Z_k + b_k k\pi) \frac{\sqrt{2} \sin(k\pi t)}{k\pi} \quad (5.18)$$

$$\text{mit } b_k = \int_0^1 b(t) \varphi_k(t) dt.$$

5.4 Anhang IV: Spezielle Grenzwertsätze

Der folgende Satz ist ein Spezialfall des Grenzwertsatzes von Lindeberg-Feller, siehe z. B. Witting und Müller-Funk (1995).

Satz 5.11 Für jedes $n = 1, 2, \dots$ seien $Z_{n,j}, j = 1, \dots, n$ unabhängige und identisch verteilte Zufallsvariablen. Es gelte $E Z_{n,j} = 0, V(Z_{n,j}) = b_n^2 \rightarrow b^2 > 0$ und

$$\lim_{c \rightarrow \infty} \sup_n E I_{[c, \infty)}(|Z_{n,1}|) Z_{n,1}^2 = 0. \quad (5.19)$$

Dann gilt

$$\mathcal{L} \left(\frac{1}{\sqrt{n}} \sum_{j=1}^n Z_{n,j} \right) \Rightarrow \mathbf{N}_{0,b^2}.$$

Wir benötigen eine Variante des Gesetzes der großen Zahlen.

Lemma 5.12 Für jedes n seien $Z_{n,1}, \dots, Z_{n,n}$ unabhängige und identisch verteilte Zufallsvariablen mit

$$\lim_{c \rightarrow \infty} \sup_n E I_{[c, \infty)}(|Z_{n,1}|) |Z_{n,1}| = 0. \quad (5.20)$$

Es sei $\mu_n = E Z_{n,1} \rightarrow \mu$. Dann gilt

$$\frac{1}{n} \sum_{i=1}^n Z_{n,i} \xrightarrow{P} \mu.$$

Die Bedingung (5.20) ist erfüllt, falls ein $\kappa > 0$ existiert mit

$$\sup_n E |Z_{n,1}|^{1+\kappa} < \infty.$$

Beweis. Seien $\tilde{Z}_{n,i} = Z_{n,i} I_{(-c,c)}(Z_{n,i}), \mu_{n,c} = E \tilde{Z}_{n,i}$. Dann gilt

$$E \left| \frac{1}{n} \sum_{i=1}^n \tilde{Z}_{n,i} - \mu_{n,c} \right| \leq \left[E \left(\frac{1}{n} \sum_{i=1}^n \tilde{Z}_{n,i} - \mu_{n,c} \right)^2 \right]^{1/2} \leq \frac{c}{\sqrt{n}}.$$

Hieraus folgt

$$\begin{aligned} E \left| \frac{1}{n} \sum_{i=1}^n Z_{n,i} - \mu \right| &\leq E |Z_{n,1} - \tilde{Z}_{n,1}| + E \left| \frac{1}{n} \sum_{i=1}^n \tilde{Z}_{n,i} - \mu_{n,c} \right| + |\mu_n - \mu_{n,c}| \\ &\leq 2 \sup_n E I_{[c, \infty)}(|Z_{n,1}|) |Z_{n,1}| + \frac{c}{\sqrt{n}} + |\mu_n - \mu|. \end{aligned}$$

Lassen wir erst n und dann c gegen Unendlich streben, so folgt

$$E \left| \frac{1}{n} \sum_{i=1}^n Z_{n,i} - \mu \right| \rightarrow 0$$

und damit die Behauptung. Die zweite Behauptung folgt aus der Ungleichung

$$\sup_n E I_{[c,\infty)}(|Z_{n,1}|) |Z_{n,1}| < \frac{1}{c^\kappa} \sup_n E |Z_{n,1}|^{1+\kappa}.$$

■

5.5 Anhang V: Quadratische Formen von normalverteilten Zufallsvariablen

In diesem Abschnitt stellen wir bekannte Aussagen über quadratische Formen von normalverteilten Zufallsvariablen in einer für unsere Zwecke geeigneten Form zusammen. Es sei X ein d -dimensionaler normalverteilter Zufallsvektor mit Erwartungswertvektor μ und der nichtsingulären Kovarianzmatrix Γ , die die Eigenwerte $\kappa_1 \geq \kappa_2 \geq \dots \geq \kappa_d$ und die orthonormierten Eigenvektoren e_1, \dots, e_d besitzt.

Lemma 5.13 *Es gilt*

$$\mathcal{L}(X^T X) = \mathcal{L} \left(\sum_{i=1}^n \kappa_i (Z_i + \kappa_i^{-1/2} e_i^T \mu)^2 \right),$$

wobei die Z_1, \dots, Z_d unabhängig und standardnormal sind.

Beweis. Sei O die orthogonale Matrix, die die Vektoren e_1, \dots, e_d als Spaltenvektoren besitzt. Dann gilt

$$\Gamma O = \Gamma(e_1, \dots, e_d) = (\kappa_1 e_1, \dots, \kappa_d e_d) = O \Lambda,$$

wobei Λ die Diagonalmatrix mit den Einträgen $\kappa_1, \dots, \kappa_d$ ist. Dann folgt

$$O^T \Gamma O = \Lambda.$$

Sei $Y = (Y_1, \dots, Y_d)^T = O^T X$. Dann hat Y die Kovarianzmatrix Λ und den Erwartungswertvektor $O^T \mu$. Dann hat Y_i den Erwartungswert $e_i^T \mu$ und die Varianz κ_i . Dann sind die $Z_i = \kappa_i^{-1/2} (Y_i - e_i^T \mu)$ standardnormal. Wegen $Y^T Y = X^T X$ folgt

$$\mathcal{L}(X^T X) = \mathcal{L} \left(\sum_{i=1}^n Y_i^2 \right) = \mathcal{L} \left(\sum_{i=1}^n \kappa_i (Z_i + \kappa_i^{-1/2} e_i^T \mu)^2 \right).$$

■

Lemma 5.14 *Es gilt*

$$\mathcal{L}(X^T \Gamma^{-1} X) = H_{d, \delta^2},$$

wobei H_{d, δ^2} die nichtzentrale Chiquadratverteilung mit d Freiheitsgraden und dem Nichtzentralitätsparameter $\delta^2 = \mu^T \Gamma^{-1} \mu$ ist.

Beweis. Es gilt $\Gamma^{-1} = O\Lambda^{-1}O^T$ und

$$\begin{aligned} X^T \Gamma^{-1} X &= Y^T O^T O \Lambda^{-1} O^T O Y = Y^T \Lambda^{-1} Y \\ &= \sum_{i=1}^n \kappa_i^{-1} Y_i^2 = \sum_{i=1}^n \kappa_i^{-1} (\kappa_i^{1/2} Z_i + e_i^T \mu)^2 \\ &= \sum_{i=1}^n (Z_i + \kappa_i^{-1/2} e_i^T \mu)^2. \end{aligned}$$

Die rechte Seite besitzt nach der Definition der nichtzentralen Chiquadratverteilung eine Chiquadratverteilung mit d Freiheitsgraden und dem Nichtzentralitätsparameter

$$\begin{aligned} \delta^2 &= \sum_{i=1}^n \kappa_i^{-1} (e_i^T \mu)^2 \\ &= (O^T \mu)^T \Lambda^{-1} (O^T \mu) = \mu^T O \Lambda^{-1} O^T \mu = \mu^T \Gamma^{-1} \mu. \end{aligned}$$

■

6 Literaturverzeichnis

- Anderson, T. W. (1984) An Introduction to Multivariate Statistical Analysis. 2nd edition, Wiley, New York.
- Behnen, K. und Neuhaus, G. (1995) Grundkurs Stochastik. 3. Auflage, Teubner Stugart.
- Billingsley, P. (1986) Covergence of Probability Measures. Wiley, New York.
- Bowman, K. O. und Shenton, L. R. (1975) Omnibus test contours for departures from normality based on $\sqrt{b_1}$ and $\sqrt{b_2}$. Biometrika 62, 243-250.
- Birkholz, H. (2012) Dissertation: Mathematical Methods for the Quantification of Actin-Filaments in Micoscopic Images.
- Cambanis, S., Huang, S. und Simons, G. (1981) On the theory of elliptically contoured distributions.
- Consul, P. C. (1966) On the exact distributions of likelihood ratio criteria for testing linear hypotheses about regression coefficients. Ann. Math. Statist. 37, 1319-1330.
- Consul, P. C. (1967) On the Exact Distributions of the Criterion W for Testing Sphericity in a p -Variate Normal Distribution. Ann. of Math. Statist. 38, 1170-1174.
- D'Agostino, R. B. und Stephens, M. A. (1986) Goodness-of-fit Techniques. Marcel Decker, New York and Basel.
- Durbin, J. (1973a) Distribution theory for tests based on the sample distribution function. Regional conference Series in Applied Mathematics 9, SIAM, Philadelphia, Pennsylvania.
- Durbin, J. (1973b) Weak convergence of the sample distribution function when parameters are estimated. Ann. Statist. 1, 279-290.
- Fang, K-T., Kotz, S. und Ng, K. W. (1990) Symmetric Multivariate and Related Distributions. Chapman and Hall, New York, London.
- Fang, K-T. und Zhang, Y-T. (1990) Generalized Multivariate Analysis. Science Press Bejing and Springer-Verlag Heidelberg.
- Farell, R. H. (1985) Multivariate Calculation (Use of the Continuous Groups). Springer-Verlag, New York, Berlin, Heidelberg, Tokyo.
- Fisher, N. I., Lewis, T. und Embleton, B. J. J. (1993) Statistical analysis of spherical data. Cambridge University Press.
- Genz, M. und Haeusler, E. (2006) Empirical processes with estimated parameters under auxiliary information. Journal of Computational and Applied Mathematics 186, 191-216.
- Girshick, M. A. (1941) The distribution of the ellipticity statistic L_e when the hypothesis is false. Terr. Magn. Atmos. Elect. 46, 455-457.
- Haywood, J. und Khmaladze, E. (2008) On distribution-free goodness-of-fit tetsting of exponentiality. Journ. Econometrics 143, 5-18.
- Hoffmann-Jørgensen, J. (1994) Probabilty With a View Towards Statistics. Vol. I, II, Chapman & Hall, New York, London.
- Jammalamadaka, S. R., SenGupta, A. (2001) Topics in Circular Statistics. World Scientific, Singapore, New Jersey, London, Hong Kong.
- Janssen, A. (2004) Asymptotic relative efficiency of tests at the boundary of regular statistical modells. Journal of Statistical Planning and Inference 126, 461-477.
- Kallenberg, O. (1997) Foundations of Modern Probability. Springer Verl. New York, Berlin.

- Khmaladze, E. (1981) Martingale approach in the theory of goodness of fit tests. *Theory of Probability and its Applications* 24(2), 283-302.
- Liese, F. und Liu, B. (2011) Randomized Goodness of Fit Tests. *Kybernetika*, 47 (6), pp. 814-839.
- Liese, F. und Miescke, K. J. (2008) *Statistical Decision Theory, Estimation, Testing, and Selection*. Springer Verl. New York, Berlin.
- Mathai, A. M. (1982) Storage capacity of a dam with Gamma type inputs. *Annals of Institute of Statistics and Mathematics* 34, 591-597.
- Mauchly, J. W. (1940a) A significance test for ellipticity in the harmonic dial. *Terr. Magn. Atmos. Elect.* 45 145-148.
- Mauchly, J. W. (1940b) Significance test for sphericity of a normal n -variate distribution. *Ann. Math. Statist.* 11 204-209.
- Matsumoto, M. und Nishimura, T. (1998) Mersenne-Twister: A 623-dimensionally equidistributed uniform pseudo-random generator. *ACM Transactions on Modeling and Computer Simulations* 8(19), 3-30.
- Moschopoulos, P. G. (1985) The distribution of the sum of independent gamma random variables. *Ann. Inst. Staist. Math.* 37, Part A, 541-544.
- Muirhead, R. (1982) *Aspects of Multivariate Analysis*. Wiley, New York.
- Neuhaus, G., (1976) Asymptotic power properties of the Cramer-von Mises test under contiguous alternatives. *Journal of Multiv. Analysis and Appl.* 6,1, 95-110.
- Pollard, D. (1984) *Convergence of Stochastic Processes*. Springer - Verlag, New York.
- Rényi, A. (1966) *Wahrscheinlichkeitsrechnung (Mit einem Anhang über Informationstheorie)*. VEB Deutscher Verlag der Wissenschaften, Berlin.
- Richter, W. D. (1991) Eine geometrische Methode in der Stochastik. *Rostocker Mathematisches Kolloquium* 44, 63-72.
- Richter, W. D. (1995) A geometric approach to the Gaussian law. In V. Mammitzsch and H. Schneeweiß, editors, *Symposia Gaussiana, Conference B: Statistical Sciences*, 25-45. de Gruyter.
- Richter, W. D. (2007) Generalized spherical and simplicial coordinates. *Journal of Multiv. Analysis and Applic.* 336, 1187-1202.
- Richter, W. D. (2009) Continuous $I_{n,p}$ -symmetric distributions. *Lithuanian Math. Journal* 49, 93-108.
- Shapiro, S. S und Francia, R. S. (1972) Approximate analysis of variance test for normality. *J. Amer. Statist. Assoc.* 67, 215-216.
- Sim, C. H. (1992) Point processes with correlated Gamma interarrival times. *Statistics & Probability Letters* 15, 135-141.
- Shorack, G. und Wellner, J. A. (1986) *Empirical Processes with Applications to Statistics*.
- Song, D. und Gupta, A. K. (1997) L_p -norm uniform distributions.
- Stephens, M. A. (1974) EDF statistics for goodness of fit and some comparisons. *J. Am. Statist. Assoc.* 69, 730-737.
- Stephens, M. A. (1976) Goodness of fit for the extreme value distribution. *Biometrika* 64, 583-588.

- Stute, W., Manteiga, W. G. und Quindimil, M. P. (1993) Bootstrap Based Goodness-Of-Fit-Tests. *Metrika* 40, 243-256.
- van der Vaart, A. W. und Wellner, J. A. (1996) *Weak convergence and Empirical Processes With Applications to Statistics*. 2nd pr. 2000. Springer, New York.
- van der Vaart, A. W. (1998) *Asymptotic Statistics*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge Univ. Press.
- Wishart, J. (1928) The generalized product moment distribution in samples from a normal multivariate population. *Biometrika* 20A.
- Witting, H. und Müller-Funk, U. (1995) *Asymptotische Statistik : parametrische Modelle und nichtparametrische Funktionale mit zahlreichen Beispielen und Aufgaben*. Stuttgart: Teubner.

7 Zusammenfassung

Die Wechselwirkung zwischen Zellen und Biomaterialien beeinflusst wesentlich die in den Zellen ausgebildete Filamentstruktur. Im ersten Teil der Arbeit werden statistische Tests zum Nachweis der Existenz einer Hauptwachstumsrichtung der Filamente konstruiert. Für normalverteilte Grundgesamtheiten sind dies bekannte Tests aus der Hauptkomponentenanalyse. Für die wesentlich größere Klasse der bivariaten homogenen Verteilungen werden neue Tests in Form bedingter Tests konstruiert, mit Computersimulationen untersucht und mit den asymptotischen Tests verglichen. Beide Typen von Tests werden auf experimentelle Daten angewandt. Im zweiten Teil der Arbeit werden mit dem Ziel der Modellwahl durch eine neuen Randomisierungstechnik Anpassungstests mit geschätzten Parametern konstruiert. Für den Cramer-von-Mises-Tests wird die explizite Form der asymptotischen Güte unter lokalen Alternativen angegeben.

8 Selbständigkeitserklärung

Hiermit erkläre ich, die vorliegende Arbeit selbständig und nur unter Verwendung der angegebenen Hilfsmittel und Quellen angefertigt zu haben.

Rostock, den 30. Mai 2012

Bing Liu